

# APPLICATIONS OF QCD EFFECTIVE THEORIES TO THE PHYSICS OF JETS AND QUARKONIUM PRODUCTION

by

**Lin Dai**

Bachelor of Engineering, Central South University, 2009

Master of Science, Nankai University, 2012

Master of Science, University of Pittsburgh, 2014

Submitted to the Graduate Faculty of  
the Kenneth P. Dietrich School of Arts and Sciences in partial  
fulfillment

of the requirements for the degree of

**Doctor of Philosophy**

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH  
KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Lin Dai

It was defended on

April 6th, 2018

and approved by

Adam Leibovich, Professor, University of Pittsburgh

Joseph Boudreau, Professor, University of Pittsburgh

Tao Han, Professor, University of Pittsburgh

Colin Morningstar, Professor, Carnegie Mellon University

Andrew Zentner, Professor, University of Pittsburgh

Dissertation Director: Adam Leibovich, Professor, University of Pittsburgh

# APPLICATIONS OF QCD EFFECTIVE THEORIES TO THE PHYSICS OF JETS AND QUARKONIUM PRODUCTION

Lin Dai, PhD

University of Pittsburgh, 2018

We apply QCD effective theories to study the physics of jets and quarkonium production. The thesis contains work in the following two related directions.

The first direction is jet physics (chapter 3 of the thesis). We introduced a function called fragmentation function to a jet (FFJ) to describe inclusive jet production from a parton and studied FFJs in different phase space and momentum regions. One of the limits we investigated was where the jet radius  $r$  was small, which lead to large  $\ln(r)$  corrections that need to be resummed. Another limit was the large  $z$  limit, where  $z$  was defined to be the fraction of energy carried by the jet from the mother parton. Here  $\ln(1 - z)$  can appear and is due to the gluon radiations that are both collinear and soft (described by collinear-soft fields). We formulated factorization theorems and used renormalization group techniques to deal with these types of logarithms. Phenomenologically, both the small jet radius and large  $z$  limit are important for comparing precision calculations with experimental data.

The second direction is quarkonium production in jets (chapter 4 of the thesis). This direction naturally combines jet physics with quarkonium production. Since quarkonium production in a jet is inclusive in the jet, we can still use inclusive quarkonium fragmentation functions that are perturbatively calculable based on the Non-relativistic QCD (NRQCD) factorization formalism. We did both analytic calculations and Monte Carlo simulations and compared them with the recent LHCb measurement of  $J/\psi$  production in jets. We found that currently existing event generators are not sufficient to study quarkonium production in jets and we proposed modifications (e.g., to PYTHIA). Both our analytic calculations

and the modified PYTHIA agreed reasonably well with the LHCb data. Our study indicates that the detailed dynamics of quarkonium production in jets can help us better understand quarkonium production mechanisms. We also studied other observables related to quarkonium production in jets that could have the potential power to clarify quarkonium production mechanisms.

## TABLE OF CONTENTS

|  |     |
|--|-----|
| <b>PREFACE</b> . . . . .   | xiv |
| <b>1.0 INTRODUCTION</b> . . . . .  | 1   |
| 1.1 An Overview of the Strong Interaction and QCD Effective Theories . . . . | 1   |
| 1.1.1 The Strong Interaction . . . . .                                       | 1   |
| 1.1.2 QCD Effective Theories . . . . .                                       | 3   |
| 1.2 Quantum Chromodynamics . . . . .   | 5   |
| 1.3 The Physics of Jets . . . . .  | 7   |
| 1.4 Quarkonium Production . . . . .  | 10  |
| <b>2.0 QCD EFFECTIVE THEORY</b> . . . . .                                    | 12  |
| 2.1 Soft-Collinear Effective Theories . . . . .                              | 12  |
| 2.1.1 Infrared divergences . . . . .   | 12  |
| 2.1.2 SCET for $\phi^3$ theories . . . . .                                   | 15  |
| 2.1.2.1 SCET Lagrangian for $\phi^3$ theory . . . . .                        | 16  |
| 2.1.2.2 Factorization of the Sudakov Factor . . . . .                        | 19  |
| 2.1.3 SCET Lagrangian for QCD . . . . .                                      | 22  |
| 2.1.4 Gauge Transformations . . . . .  | 26  |
| 2.1.5 Wilson Lines . . . . .   | 29  |
| 2.1.6 Reparameterization Invariance . . . . .                                | 34  |
| 2.1.7 Factorization Theorems . . . . .                                       | 35  |
| 2.2 Non-Relativistic QCD . . . . .   | 38  |
| 2.2.1 NRQCD Power Counting Parameter $v$ . . . . .                           | 38  |
| 2.2.2 NRQCD Lagrangian . . . . .   | 39  |

|            |  |           |
|------------|--|-----------|
| 2.2.3      | Quarkonium Production and NRQCD Factorization Formalism . . .                      | 42        |
| <b>3.0</b> | <b>FRAGMENTATION TO A JET . . . . .</b>  | <b>45</b> |
| 3.1        | Fragmentation of a Jet with Small Radius . . . . .                                 | 45        |
| 3.1.1      | Introduction . . . . .   | 45        |
| 3.1.2      | Inclusive Jet Fragmentation Function . . . . .                                     | 47        |
| 3.1.2.1    | NLO Calculation of Quark FFJ . . . . .   | 50        |
| 3.1.2.2    | NLO Calculation of Gluon FFJ . . . . .   | 55        |
| 3.1.2.3    | Renormalization Scaling Behavior . . . . .   | 57        |
| 3.1.3      | Factorization Theorem for the Fragmentation inside a Jet . . . . .                 | 58        |
| 3.1.4      | Subjet Fragmentation inside a Fat jet . . . . .                                    | 63        |
| 3.1.5      | Conclusions . . . . .  | 68        |
| 3.2        | Fragmentation to a jet in the large $z$ limit . . . . .                            | 69        |
| 3.2.1      | Introduction . . . . .   | 69        |
| 3.2.2      | The FFJ in the limit $z \rightarrow 1$ . . . . .                                   | 72        |
| 3.2.2.1    | Factorization of the FFJ when $z \rightarrow 1$ . . . . .                          | 74        |
| 3.2.2.2    | NLO calculation of the FFJ near the endpoint . . . . .                             | 78        |
| 3.2.3      | Renormalization Group Evolution and Resummation of Large Logarithms . . . . .      | 81        |
| 3.2.3.1    | RG evolution from the factorization of the FFJ . . . . .                           | 81        |
| 3.2.3.2    | Contribution of nonglobal logarithms . . . . .                                     | 84        |
| 3.2.4      | Numerical Results . . . . .  | 85        |
| 3.2.5      | Conclusion and Outlook . . . . .   | 88        |
| <b>4.0</b> | <b>STUDIES OF QUARKONIUM PRODUCTION IN A JET . . . . .</b>                         | <b>91</b> |
| 4.1        | Analytic and Monte Carlo Studies of Jets with Heavy Mesons and Quarkonia . . . . . | 91        |
| 4.1.1      | Introduction . . . . .   | 91        |
| 4.1.2      | Fragmenting Jet Functions with Angularities . . . . .                              | 95        |
| 4.1.3      | $e^+e^- \rightarrow 2$ Jets with a $B$ Meson . . . . .                             | 98        |
| 4.1.4      | $e^+e^- \rightarrow 3$ Jets with the Gluon Jet Fragmenting to $J/\psi$ . . . . .   | 104       |
| 4.1.5      | Conclusion . . . . .   | 111       |
| 4.2        | NRQCD Confronts LHCb Data on Quarkonium Production within Jets . . . . .           | 113       |

|              |   |     |
|--------------|---|-----|
| 4.3          | Quarkonium Polarization and the Long Distance Matrix Elements Hierar- |     |
|              | chies using Jet Substructure . . . . .                                | 121 |
| 4.3.1        | Introduction . . . . .  | 121 |
| 4.3.2        | The Fragmenting Jet Functions . . . . .                               | 123 |
| 4.3.3        | Discussion of the $J/\psi$ production mechanisms . . . . .            | 125 |
|              | 4.3.3.1 Normalized $J/\psi$ production cross section . . . . .        | 125 |
|              | 4.3.3.2 Normalization using 1-jet inclusive cross section . . . . .   | 128 |
| 4.3.4        | Conclusion . . . . .  | 131 |
| 5.0          | SUMMARY . . . . .   | 132 |
| APPENDIX A.  | . . . . .   | 134 |
| A.1          | Hadron Fragmentation inside a Jet . . . . .                           | 134 |
| APPENDIX B.  | . . . . .   | 139 |
| B.1          | Renormalization Group and Resummation . . . . .                       | 139 |
|              | B.1.1 Evolution of Measured and Unmeasured Functions . . . . .        | 139 |
|              | B.1.2 Plus-distribution identities . . . . .                          | 141 |
|              | B.1.3 Reorganization of logarithms of $(1 - z)$ . . . . .             | 142 |
|              | B.1.4 Profile Functions . . . . .                                     | 143 |
| B.2          | Matching Coefficients and Consistency Checks . . . . .                | 145 |
|              | B.2.1 Evaluation of matching coefficients . . . . .                   | 145 |
|              | B.2.2 Sum Rules . . . . .   | 149 |
| APPENDIX C.  | . . . . .   | 150 |
| C.1          | Unnormalized and Normalized cross sections for Bodwin et.al . . . . . | 150 |
| C.2          | Unnormalized and Normalized cross sections for Chao et.al . . . . .   | 152 |
| C.3          | Normalization using only color octet channels . . . . .               | 153 |
| C.4          | Insensitivity to $z_{min}$ and $z_{max}$ . . . . .                    | 154 |
| C.5          | Lower $z$ plots . . . . .   | 155 |
| BIBLIOGRAPHY | . . . . .   | 156 |

## LIST OF TABLES

|   |   |     |
|---|---|-----|
| 1 | Three types of reparametrization invariance in SCET . . . . .   | 34  |
| 2 | Quarkonium mass and momentum scales [1]. . . . .  | 39  |
| 3 | Power counting rules for basic NRQCD operator building blocks [1]. . . . .                                      | 41  |
| 4 | Characteristic scales of the different functions in the factorization theorem of<br>Eq. (4.16). . . . .         | 103 |
| 5 | LDMEs for NRQCD production mechanisms. We use central values taken from<br>global fits in Refs. [2, 3]. . . . . | 105 |
| 6 | LDMEs for NRQCD production mechanisms used in this section in units of<br>$\text{GeV}^3$ . . . . .              | 115 |



## LIST OF FIGURES

|    |  |    |
|----|--|----|
| 1  | Feynman diagrams for gluon self-interactions. . . . .  | 7  |
| 2  | The dependence of strong coupling constant $\alpha_s$ on energy scale. The plot is quoted from Particle Data Group (PDG) publication Ref. [4]. . . . .   | 8  |
| 3  | Infrared divergence in $\phi^3$ theory. One of the external particles with momentum $p$ splits into two with momenta $p_1$ and $p_2$ . . . . .   | 13 |
| 4  | Virtual correction to $\mathcal{M}_n$ needed to cancel infrared divergences that appear in $\mathcal{M}_{n+1}$ . . . . .   | 15 |
| 5  | Sudakov form factor for $\phi^3$ theory. . . . .   | 20 |
| 6  | Radiation of an ultra-soft gluon from a collinear quark. . . . .   | 31 |
| 7  | Radiation of two gluons from a collinear quark. . . . .  | 32 |
| 8  | Sudakov form factor for QCD. . . . .   | 36 |
| 9  | Jet fragmentation at NLO in $\alpha_s$ . Diagram (a) shows the jet merging, so the contribution to FFJ should be proportional to $\delta(1 - z)$ . Diagram (b) shows the jet splitting, which has a contribution with a fraction $z < 1$ . . . . . | 49 |
| 10 | Feynman diagrams for quark jet splitting contribution at NLO in $\alpha_s$ . Here the dashed lines represent the unitary cuts. The gluon in the final state is outside the jet. Diagram (a) has its Hermitian conjugate contribution. . . . .      | 53 |
| 11 | Feynman diagrams for gluon jet splitting contribution at NLO in $\alpha_s$ . Here the dashed lines represent the unitary cuts. The quark in the final state is outside the jet. Diagram (a) has its Hermitian conjugate contribution. . . . .      | 55 |

|    |  |    |
|----|--|----|
| 12 | Feynman diagrams for jet splitting contributions to jet fragmentation initiated by gluon parton. Diagram (a) has its Hermitian conjugate contribution. Diagram (a) and (b) represents $g \rightarrow J_g$ splitting, and Diagram (c) for $g \rightarrow J_q$ splitting. . . . .  | 56 |
| 13 | Fragmentation process from the parton ( $p$ ) to the hadron ( $p_H$ ) through the jet ( $p_J$ ). . . . .   | 60 |
| 14 | Feynman diagrams of real gluon emissions for the subjet quark fragmentation inside a jet at NLO. Diagram (a) has its Hermitian conjugate contribution. .   | 64 |
| 15 | Phase space for the real gluon emission in the soft function. In $(k_+, k_-)$ plane the region above the border line $k_- = t^2 k_+$ gives out-jet contribution and the region in the below gives in-jet contribution. $\Lambda_+$ is the maximum value for the distribution of $\ell_+$ and can be chosen arbitrarily. . . . .  | 79 |
| 16 | $D_{J_q/q}(z)$ (left panel) and $D_{J_g/g}(z)$ (right panel) with different jet energies. Red, blue, and black curves correspond to jet energy $E_J$ equal to 500, 1000, and 2000 GeV, respectively. Jet radius is chosen as $R = 0.2$ and the factorization scale is set as $\mu_f = E_J$ . Error estimation is described in the text. . . . .  | 85 |
| 17 | $D_{J_q/q}(z)$ (left panel) and $D_{J_g/g}(z)$ (right panel) with different jet radii. Red, blue, and black curves correspond to jet radius $R$ equal to 0.1, 0.2, and 0.4 respectively. Jet energy are given as $E_J = 1000$ GeV and the factorization scale is set as $\mu_f = E_J$ . . . . .  | 86 |
| 18 | Comparison of the result using leading DGLAP evolution and the resummed result at $\text{NLL}_G + \text{NLO}$ from the factorization approach. The orange (green) curves are obtained using leading DGLAP evolution with FFJs running from $\mu_c = E_J R$ ( $\mu_{cs} = E_J R(1 - z)$ ) to $\mu_f = E_J$ . Blue curves are the resummed result of the FFJs. $R = 0.2$ , $E_J = 1000$ GeV. . . . . | 87 |
| 19 | Comparison of the resummed results with (blue) and without resumming the NGLs (orange). Here $R = 0.2$ and $E_J = 1000$ GeV. . . . .   | 88 |
| 20 | Profile function $\mu_{cs,PF}$ (solid black curve and gray band, defined in Eq. (3.122)) is used to estimate errors due to variation of the collinear-soft scale. The dashed line is the $z$ dependent collinear-soft scale $\mu_c(1 - z)$ . $\mu_c = 200$ GeV. .  | 89 |

|    |  |     |
|----|--|-----|
| 21 | The $z$ distributions for $d\sigma(\tau_0, z)$ at $\tau_0 = (1.5, 2.0, 2.5) \times 10^{-3}$ for analytic calculations with theoretical uncertainty are shown in green. Monte Carlo simulations using Madgraph + PYTHIA and Madgraph + Herwig are shown in black and red, respectively. . . . .                           | 101 |
| 22 | Analytic results for the $z$ distributions of $d\sigma(\tau_0, z)$ at $\tau_0 = (1.5, 2.0, 2.5) \times 10^{-3}$ . The orange curve is calculated with a measured jet scale that does not depend on $z$ whereas the green curve uses a scale that does depend on $z$ (as in Fig. 21).                                     | 102 |
| 23 | Angular distributions of $d\sigma(\tau_a, z)$ for $a = 0$ at $z = 0.4, 0.6, 0.8$ . Analytic results are shown as green bands. Monte Carlo results are shown as black lines for Madgraph + PYTHIA and red lines for Madgraph + HERWIG. . . . .  | 103 |
| 24 | Angular distributions of $d\sigma(\tau_a, z)$ for $a = 0$ at $z = 0.1, 0.3, 0.5, 0.7$ . Analytic calculations are shown as red (green) bands for the $^3S_1^{(8)}$ ( $^1S_0^{(8)}$ ) production mechanisms. Results from Madgraph + PYTHIA are shown as red (green) dashed lines for the same mechanisms. . . . .        | 108 |
| 25 | Angular distributions of $d\sigma(\tau_a, z)$ for $a = +1/2, 0, -1/2, -1$ at $z = 0.5$ . Analytic calculations are shown as red (green) solid lines for the $^3S_1^{(8)}$ ( $^1S_0^{(8)}$ ) production mechanisms. Results from Madgraph + PYTHIA are shown as red (green) dashed lines for the same mechanisms. . . . . | 109 |
| 26 | $z$ distributions of $d\sigma(\tau_a, z)$ for NLL' analytic calculations (bands), PYTHIA (dashed lines), and GFIP (solid lines) for fixed values of $\tau_0 = (4, 5, 6) \times 10^{-3}$ . . . . .  | 110 |
| 27 | PYTHIA predictions for $c$ quark and gluon $z$ distributions (where $z$ is the fraction of the energy of the parton initiating the jet) after showering to the scale $2m_c$ . . . . .  | 118 |
| 28 | Predicted $z(J/\psi)$ distribution using GFIP (gray) and FJF (red) for the three choices of LDME in Table 1 and the LHCb measurements of $z(J/\psi)$ . . . . .   | 119 |
| 29 | Cross sections for inclusive gluon and charm jets at the LHC. The center of mass energy is $\sqrt{s} = 13$ TeV. . . . .  | 125 |

|    |  |     |
|----|--|-----|
| 30 | Cross sections for the different production channels at $z = 0.4, 0.5$ , and $0.6$ as a function of the jet energy. The first two rows show the unnormalized cross sections ( $d\sigma_i/dEdz$ ), with the second row showing plots normalized to unit area for a better visualization of the shapes, i.e., we multiply each curve of the first row by an appropriate constant to get the corresponding curve in the second row. Similar plots for the normalized cross section ( $d\tilde{\sigma}_i/dEdz$ ) are shown in the third and fourth row. The LDMEs are from Butenschoen et al.'s extractions [2]. . . . . | 127 |
| 31 | Total normalized cross section (i.e. $d\tilde{\sigma}/dEdz$ defined in Eq. (4.34)) with error bands. Red, black, and blue curves correspond to Bodwin et al. [5], Butenschoen et al. [2, 3], and Chao et al.'s [6] extractions, respectively. . . . .  | 128 |
| 32 | Total normalized cross section (i.e. $d\hat{\sigma}/dEdz$ defined in Eq. (4.36)) with error bands. Red, black, and blue curves correspond to Bodwin et al. [5], Butenschoen et al. [2, 3], and Chao et al.'s [6] extractions, respectively. . . . .  | 129 |
| 33 | Comparisons of the production channels for various LDMEs using Eq. (4.35). Last row shows the plots normalized to unit area. This is indicated by $1/\hat{\sigma}_i$ for the cross section label in the fourth row, which also cancels the LDME dependence of the numerator. . . . .   | 130 |
| 34 | Feynman diagrams of real gluon emissions for quark fragmentation inside a jet at NLO. The gluon in the final state is also inside a jet. Diagram (a) has its Hermitian conjugate contribution. . . . .   | 135 |
| 35 | Profile functions for $\mu_S^{PF}(\tau_0)$ and $\mu_J^{PF}(\tau_0)$ , the $\tau_0$ -dependent renormalization scales that we use in the scale variations of our measured soft function and measured jet function. Also shown are traditional scale variations done by varying $\mu$ by $\pm 50\%$ . . . . .  | 143 |
| 36 | Unnormalized and normalized cross sections for Bodwin et al. extractions [5]. The conventions followed are same as in Fig. (30). . . . .   | 151 |
| 37 | Unnormalized and normalized cross sections for Chao et al. extractions [6]. The conventions followed are same as in Fig. (30). . . . .   | 152 |

|    |  |     |
|----|--|-----|
| 38 | Cross section normalized by ignoring the $^3S_1^{[1]}$ channel contribution in Eq. 4.33. The second and fourth row are obtained by normalizing the curves in the first and third row to unit area respectively. . . . .                    | 153 |
| 39 | Solid curves correspond to $(z_{min}, z_{max}) = (0.3, 0.8)$ and the dashed curves $(z_{min}, z_{max}) = (0.4, 0.7)$ . Due to the change in normalization, all the curves shift upwards without changing their qualitative shapes. . . . . | 154 |
| 40 | Lower $z$ plots for the cross section (Eq. (4.35)). The conventions followed are same as those in Fig. (33). . . . .   | 155 |

## PREFACE

The past six years in graduate school has been such a venturous and enjoyable journey. I would like to take this opportunity to express my thankfulness.

First of all, I owe so much to my Ph.D. advisor Adam Leibovich for his guidance of my graduate study, with such patience and constant encouragement. I always felt refreshed and wanted to do more physics after discussions with him. His natural talent of physics intuition and the clear way of explaining physics has always been inspiring to me.

I would like to thank my collaborators Reggie Bain, Andrew Hornig, Chul Kim, Yiannis Makris, Thomas Mehen, Prashant Shrivastava, from whom I have learned so many technical details for doing physics. I would like to thank my Ph.D. committee since each past meeting was a boost to my Ph.D. research.

I would like to thank all the students at PITT and CMU with whom I took classes and the professors (Joe Boudreau, Daniel Boyanovsky, Hael Collins, Anthony Duncan, Aryes Freitas, Richard Holman, Kaoru Hagiwara, Tae Min Hong, Adam Leibovich, Vincent Liu, Ralph Roskies, Eric Swanson, Ira Rothstein, Andrew Zentner) who taught us. They were part of the reasons why the graduate school years are so enjoyable.

I would like to thank my officemates Junmou Chen, Zhuoni Qian, Xing Wang, Daniel Wiegand, Yang Ma, Hongkai Liu, who always make me realize I am not doing physics alone after all.

I would like to thank Barmak Shams Es Haghi with whom I spent so much time learning and discussing physics and mathematics. Even though we have not yet, I believe we will write papers together on fascinating topics in the future.

I would like to thank my undergraduate mentor Yong He. His enthusiasm about physics and encouragement is part of the reasons why I wanted to pursue a career in physics.

Finally, thank you is just too little a phrase for the unconditional love and support I received from my parents during the entire time of my existence.

## 1.0 INTRODUCTION

The structure of this thesis is as follows. Chapter 1 is a brief description of Jet physics and quarkonium production that are relevant to our work. Chapter 2 is a review of the two types of QCD effective theories (soft-collinear effective theory and non-relativistic QCD) that are used in our work. Chapter 3 is devoted to the study of inclusive jet production and Chapter 4 to quarkonium production in jets.

## 1.1 AN OVERVIEW OF THE STRONG INTERACTION AND QCD EFFECTIVE THEORIES

### 1.1.1 The Strong Interaction

The study of particle physics is about the basic building blocks of the universe and their interactions. Even though it is far from being completed, the standard model of particle physics (SM for short) is our current best understanding of these building blocks and their interactions. According to the SM, there are three families of quarks, i.e.  $(u, d)$ ,  $(c, s)$ , and  $(t, b)$ , and three families of leptons, i.e.,  $(\nu_e, e)$ ,  $(\nu_\mu, \mu)$ , and  $(\nu_\tau, \tau)$ . All the quarks and leptons are fermions with spin  $1/2$ . The interactions of these fermions are mediated by spin 1 gauge bosons  $Z, W^\pm$  bosons and photons (responsible for the electro-weak interaction), and gluons  $g$  (responsible for the strong interaction). There is an additional spin 0 particle called the Higgs boson which gives masses to all the other fundamental particles (except for the neutrinos) through spontaneous symmetry breaking.

The establishment of the SM is the results of heroic journeys of generations of physicists.



Here I just mention some of the milestones that are key for the understanding of the strong interaction. If we want, we can go back to the philosophical concepts of atoms in ancient Greece. But let's just start from the experiments by Rutherford and his collaborators between 1908 and 1913. In those series of experiments,  $\alpha$  particles (made of Helium-4 nucleus, i.e., 2 protons and 2 neutrons) were bombarded into thin gold foils, the surprising thing was that there was a significant number of  $\alpha$  particles that were deflected at large angles relative to the incidence direction, this can only happen when the  $\alpha$  particles are hitting on targets with masses concentrating in small regions and the quantity of masses in those small regions are comparable to that of the  $\alpha$  particle. Those small regions of large mass concentration are called nuclei. Based on the analysis of coulomb potential interaction, Rutherford estimated the size of a typical nucleus to be around  $10^{-15}$  m [7]. It was not until 1932 that the neutron was discovered [8], which eventually established that a nucleus is made of protons and neutrons (except for rare cases like the hydrogen atom).

Whether a nucleus is made of just protons or protons plus neutrons, there is the puzzle of what forces are responsible for holding them together. After all, the electric force among protons grows fast (proportional to  $1/r^2$ ) as they get close. It must be some kind of force that is stronger than the electric force (i.e., the name strong interaction). In 1935 Yukawa suggested that the interaction among nucleons be mediated by massive bosons (called pions  $\pi$ 's) [9]. The basic idea is that the force due to pions are short ranged (length of the range is around  $1/m_\pi$ ) with potential that looks like

$$V_{\text{Yukawa}}(r) \propto \frac{e^{-m_\pi r}}{r}, \quad (1.1)$$

in contrast to the well-known long-ranged coulomb potential

$$V_{\text{Coulomb}}(r) \propto \frac{1}{r}, \quad (1.2)$$

which is the result of massless photon exchanging among charged particles and which can be thought of as the  $m_\pi \rightarrow 0$  limit of the Yukawa limit. Between 1930's and 1960's, there were lots of efforts following the framework proposed by Yukawa without going much further.

Right after the discovery of the neutron, Heisenberg proposed a symmetry ( $SU(2)$  symmetry, or isospin symmetry) between the proton and the neutron in a nucleus, mostly because

the masses of the proton and the neutron are very close. To put this direction further, in the 1960s, Gell-Mann categorized the discovered (and predicted to exist) mesons and baryons similar to the periodic table of Mendeleev. The “periodic table of Gell-Mann” are based on quark models in which particles are made of  $u$ ,  $d$ ,  $s$  quarks and the three quarks satisfy  $SU(3)$  symmetry. The quark model seemed to be in violation with the spin-statistics theorem (particles with integer spin are bosons and those with half odd integer spin are fermions). The existence of a particle called  $\Omega^-$  whose quark content is three  $s$  quarks ( $sss$ ) is such an example. Most then existing problems of the quark model disappeared if a new internal degree of freedom was introduced, which Gell-Mann coined the name “color” for the new internal degree of freedom.

The exploration of symmetry groups eventually lead to the establishment of Non-Abelian gauge theories. Based on the procedure of renormalization to remove infinities, quantum field theory had been very successful in comparing Quantum Electrodynamics (QED) calculations and experimental data. The interaction of QED is based on  $U(1)$  gauge symmetry, which says the physics (or the classical equation of motion) is independent of a local phase transformation. In a 1954 paper by Yang and Mills, they generalized  $U(1)$  gauge interaction of QED to Non-Abelian gauge groups [10]. Eventually, Non-Abelian Gauge theories became the foundations of the SM, including the Quantum Chromodynamics (QCD) based on the  $SU(3)_c$  gauge symmetry of the color degree of freedom. Initially, the application of QCD was not successful to the extent that people were ready to give up on quantum field theory in describing the strong interaction and focused on objects such as scattering amplitude methods which rely solely on basic principles (such as causality, unitarity, analyticity). Only after the establishment of asymptotic freedom was the  $SU(3)_c$  gauge theory widely accepted as the main tool for studying the strong interaction [11].

### 1.1.2 QCD Effective Theories

In the next section, we will lay out the technical details of QCD. This section focuses only on some motivations of using QCD effective field theories (EFTs) to study the strong interaction.

Generally speaking, every successful physical theories we already knew can be thought of

as an effective theory. For instance, Newtonian gravity is an effective theory of the General Relativity (in the weak gravitation limit). The General Relativity and The SM can be thought of as effective theories of some yet unknown theory at low energy limit. One reason of studying effective theories is that the more inclusive a theory is, the harder it tends to be for detailed applications. For instance, to calculate the light emission spectrum of the hydrogen atom, it is enough to use Non-Relativistic Schrodinger equation without having to use the full complicated machinery of QED. Another reason that effective theory methods are so useful is that there are interesting physics at all distance (or equivalently energy) scales and we can do physics reliably in different scales (either length or energy scales). For instance, condensed matter physicists do not need calculate the dynamics of quarks to study properties of materials since quarks lie in nuclei with distance of order  $10^{-15}$  m while material properties are determined by atomic structures with length scale around  $10^{-10}$  m.

The motivations of studying QCD effective theories are the same as those of studying general effective theories. There are only very limited situations where we can do exact QCD calculations. Otherwise, we can only rely on numerical simulations (or mostly Monte Carlo methods). On the other hand, modern collider experiments (such as those carried out at the LHC) involve complicated multi-scale processes. EFT allows us to study different scales separately and use factorization theorems to combine all relevant components. Take a typical event at the LHC for example, two protons (with energy scale around 1 GeV) collide and leave thousands of tracks in the detectors. These tracks reflect the hardest collision structures (at energy scale around 1 TeV ), and they are made of hadrons (such as pions, kaons, at GeV scales) and leptons (electrons at MeV scale and muons around 100 MeV).

As was mentioned above, the discovery of asymptotic freedom promoted QCD as the main theory of the strong interaction. Asymptotic freedom (more details on this in the next section) means that the interaction of partons (quarks and gluons) decreases as the energy scale increases, and only in those higher energy scales can we rely on perturbative calculations. For the LHC experiments, two of the perturbatively calculable parts are: (1) The hard processes which describe the partons from the two incoming protons collide at TeV scales. (2) The splitting of partons after the hard collision before hadrons (such as pions and quarkonia) are produced, which is called parton shower. The splitting of partons tends to be

collinear, which means splitting at small relative angles (the reason is illustrated in Section 1.3 and 2.1.1). As a result, the tracks recorded by the detectors tend to be collimated along certain directions, and we call these collimated tracks jets. Note that since these tracks are hadrons, we need to combine the perturbative calculations (responsible for the parton shower) with the non-perturbative parts (responsible for the formation of hadrons). This thesis focuses on the study of jets and quarkonium production, which involves both perturbative parts and transitions to non-perturbative parts. And effective field theories enable us to analytically use QCD to study such complicated experiments that are carried out at the LHC.

## 1.2 QUANTUM CHROMODYNAMICS

Quantum Chromodynamics (QCD) is a theory of the strong interaction based on  $SU(3)$  gauge symmetry and is a basic building block of the present standard model of particle physics. It was gradually established in the 1960s and 1970s when a large number of hadrons were discovered and physicists were trying to find order in an explosive collection of experimental data.

$SU(3)$  is a Lie group in which any element  $U$  satisfies

$$U^\dagger U = UU^\dagger = 1, \quad \det(U) = 1. \quad (1.3)$$

This Lie group (being simply connected) is uniquely determined by its group generators that form a Lie algebra under the commutation relation

$$[T^a, T^b] = if^{abc}T^c, \quad (1.4)$$

i.e., any element in  $SU(3)$  can be generated by the exponentiation operation

$$e^{i\theta^a T^a} \quad (1.5)$$

with a certain tuple of real numbers  $\theta^a$ . Eq. (1.3) and (1.5) require the generators  $T^a$  to be hermitian and traceless, which leads to 8 generators for  $SU(3)$  ( $N^2 - 1$  generators for  $SU(N)$ ).

$T^a$ 's are usually chosen to have the following forms (Gell-Mann matrices)

$$\begin{aligned} \lambda^1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda^2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda^3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ \lambda^4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, & \lambda^5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, & \lambda^6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\ \lambda^7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, & \lambda^8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \end{aligned} \quad (1.6)$$

with  $T^a \equiv \lambda^a/2$ .

The Lagrangian of QCD is

$$\mathcal{L}_{QCD} = \bar{\psi}_i (i \not{D} - m)_{ij} \psi_j - \frac{1}{4} F_{\mu\nu}^a F^{a\mu\nu}, \quad (1.7)$$

where

$$[D_\mu]_{ij} \equiv \delta_{ij} i \partial_\mu - g_s A_\mu^a T_{ij}^a, \quad (1.8)$$

and the field strength tensor

$$F_{\mu\nu} \equiv \frac{i}{g_s} [D_\mu, D_\nu] \Rightarrow F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g f^{abc} A_\mu^b A_\nu^c. \quad (1.9)$$

The non-abelian nature of  $SU(3)$  gives rise to gauge field self-interactions whose Feynman diagrams include vertices with 3 and 4 gluon lines (see Figure 1),

$$-g_s f^{abc} (\partial_\mu A_\nu^a) A^{b\mu} A^{c\nu}, \quad -\frac{1}{4} g_s^2 f^{eab} f^{ecd} A_\mu^a A_\nu^b A^{c\mu} A^{d\nu}, \quad (1.10)$$

which make the infrared divergences involved in QCD much more dire than that of QED, whose gauge field photons have no self-interaction (some more details on infrared divergences are discussed in Section 2.1.1).

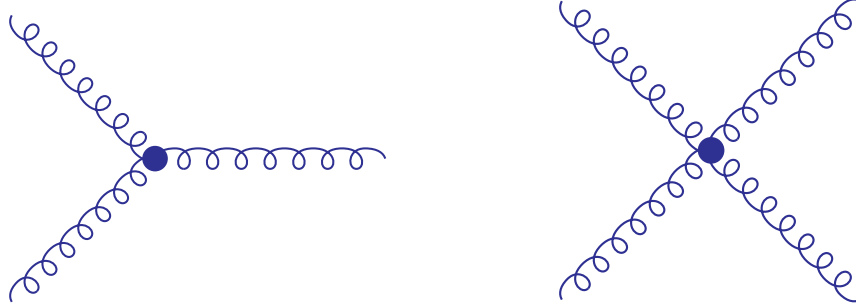


Figure 1: Feynman diagrams for gluon self-interactions.

The key feature of the  $SU(3)$  gauge theory is asymptotic freedom [11, 12], which is the basis for the validity of perturbative calculations. Figure 2 shows the running of the coupling constant  $\alpha_s$  ( $\alpha_s \equiv g^2/4\pi$ ) with respect to energy scale, which indicates that the interaction is weaker towards ultra-violet (UV) scales (the meaning of asymptotically free at UV).

### 1.3 THE PHYSICS OF JETS

Jets in high energy collisions have been an important theoretical and experimental probe of physics for decades. Currently, they are not only important for understanding Quantum Chromodynamics (QCD), but are crucial in our searches for beyond the Standard Model physics at the Large Hadron Collider at CERN and will continue to be important for any future collider that may be built. Understanding the property of jets and being able to calculate reliable cross sections to compare to data are thus extremely important to current and future studies in particle physics.

Take experiments at the LHC for example. Two protons are accelerated to extremely high energies (currently with center of mass energy  $\sqrt{s} = 13$  TeV) and collide. After the collision, lots of particles are generated and they leave in the detectors bunches of tracks. As a general feature of quantum mechanics, degeneracy of states leads to singularities. As a familiar example, the second order corrections to energy levels in time-independent perturbation

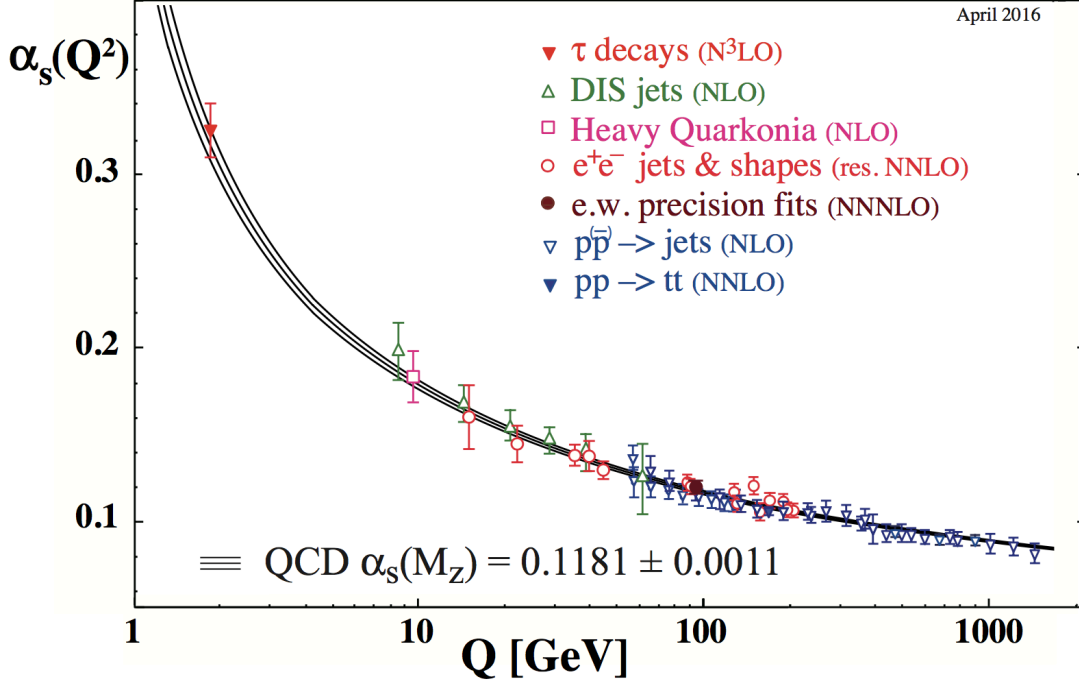


Figure 2: The dependence of strong coupling constant  $\alpha_s$  on energy scale. The plot is quoted from Particle Data Group (PDG) publication Ref. [4].

theories are:

$$\delta E_n^{(2)} = \sum_{m \neq n} \frac{|\langle m | H_I | n \rangle|^2}{E_n - E_m} \quad (1.11)$$

If  $E_n$  is degenerate, or there are lots of states with energies close to  $E_n$ , the contributions from those states are enhanced. Similarly, in jet physics, collinear splittings, which correspond to producing almost degenerate states of the initial state, are enhanced comparing to other splittings in random directions. We will discuss more on collinear singularities in Section 2.1.1. With certain algorithms, these tracks are clustered into single objects called jets. Study of these jets is essential. On the one hand, it can deepen our understanding of QCD itself. On the other hand, since jets are ubiquitous in collider experiments, if one wants

to extract possible new physics information from collider experiments, analysis of jets is unavoidable. Loosely speaking, jets are defined to be collections of particles within some solid angles. How to choose the solid angle, however, is a very subtle issue. Below we will discuss two main categories of jet algorithms to define what a jet is.

The analysis of QCD jets can be dated back to 1977 when Sterman and Weinberg managed to interpret infrared divergent cross sections for di-jet productions of  $e^+e^-$  collisions [13]. In their analysis, a di-jet event is required to have at least a fraction  $1 - \epsilon$  of energy deposited in a cone of half angle  $\delta$  along the jet axis (for a back-to-back two-jet event, the jet axis is well defined).

There are two types of jet algorithms that are used in present day collider physics [14]. The first type are called cone-jet algorithms that are similar, but not identical, to the one used in Ref. [13]. In a cone-jet algorithm (or iterative cone-jet algorithm), a seed particle is chosen first, then all particles within a radius of  $R$  are collected to form a jet. The metric for defining the radius  $R$  (which is dimensionless) is

$$d_{ij} = \sqrt{(\Delta\phi)_{ij}^2 + (\Delta\eta)_{ij}^2}, \quad (1.12)$$

where  $\phi$  is the azimuthal angle and  $\eta$  the pseudo-rapidity defined as

$$\eta \equiv \ln \left( \cot \left( \frac{\theta}{2} \right) \right), \quad (1.13)$$

with  $\theta$  the polar angle. Here both  $\phi$  and  $\theta$  are defined with respect to the collision beam axis. The second are called  $k_T$ - (or clustering-) type jet algorithms which are widely used in, for instance, LHC experiments. For this type of jet algorithms, a metric  $d_{ij}$  (which is typically different from Eq. (1.12)) is defined first and then the following algorithm is followed to define jets with radius  $R$ :

- (1) Calculate  $d_{ij}$  for all final particles, find the minimum of  $d_{ij}$  denoted as  $d_{i_m j_m}$ ;
- (2) If  $d_{i_m j_m} \leq R$ , merge the four momenta of particle  $i_m$  and  $j_m$  to form a new final particle (remove the original  $i_m$  and  $j_m$  from the final particle list);
- (3) Repeat (1) until there are no pairs whose distances are smaller than  $R$ .

The choice of a jet algorithm depends heavily on what process is under study. From a theoretical point view, some jet algorithms might be better suited than others for analysis



such as deriving factorization theorems. No matter what jet algorithms are used, the ultimate goal is to use jet structures to reconstruct the hard interactions (interactions of the hardest partons).

## 1.4 QUARKONIUM PRODUCTION

In this thesis (as is typically used in high energy community), quarkonium refers to heavy quark bound states charmonium and bottomonium. There are many aspects of quarkonium physics that are fascinating [15] such as quarkonium spectroscopy and decay. We will be focusing on quarkonium production. Historically, the discovery of charmonium ( $J/\psi$ ) was very important for the establishment of the standard model and initiated the “November Revolution” in the high energy physics community.

Quarkonium physics continues to be an active field of research. The unique feature of a quarkonium system is that it involves multiple regimes of the scales of QCD: from the heavy quark mass scale  $m_Q$ , where the strong coupling  $\alpha_s$  is still reasonably small for perturbative calculations (ref. Figure 2,  $\alpha_s(m_{\text{charm}}^2) \sim 0.35$  and  $\alpha_s(m_{\text{bottom}}^2) \sim 0.21$  if  $m_{\text{charm}} = 1.4$  GeV and  $m_{\text{bottom}} = 4.8$  GeV), to the scales which are very close to the  $\Lambda_{QCD}$  (such as  $m_Q v$ ,  $m_Q v^2$ , more on this in Section 2.2.1). So quarkonia are special systems to probe the transition of QCD from the perturbative to non-perturbative regions.

Many of the past studies of quarkonium production are based on the NRQCD factorization formalism (discussed in Section 2.2). This formalism predicts different quarkonium production mechanisms where the only inputs are a set of long distance matrix elements (LDMEs). The LDMEs are supposed to be universal, so that once measured they can be used to predict all other measurements involving those LDMEs. However, extractions of LDMEs from different data gave very different numerical values, which in turn gave quite different predictions. One of the most prominent examples is the polarization puzzle [15], where different LDME extractions gave rise to completely different predictions of polarization of quarkonium production. Continuing study of these issues will give us a better understanding of QCD in the regime where the validity of the perturbative calculations approaches its

limit.

## 2.0 QCD EFFECTIVE THEORY

In this thesis, we focus on applications of two types of effective theories of QCD, i.e., soft-collinear effective theory (SCET) and non-relativistic QCD (NRQCD), which we will be reviewing in this chapter. SCET is suitable for the descriptions of high energy jets and NRQCD is a powerful tool to study quarkonia.

### 2.1 SOFT-COLLINEAR EFFECTIVE THEORIES

SCET was invented to properly deal with QCD infrared divergences and it is a very useful framework to study multi-scale problems in modern collider experiments [16, 17, 18, 19, 20, 21]. Similar to other types of effective field theories, SCET facilitates the derivations of factorization theorems involved in multi-scale physical processes and enables resummations of large logarithms by employing renormalization group techniques.

#### 2.1.1 Infrared divergences

Consider a scattering process, as is shown in Figure 3, for the massless scalar  $\phi^3$  theory (in  $D$  dimensional space-time) with Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{g}{3!}\phi^3. \quad (2.1)$$

Let  $\mathcal{M}_n$  be the amplitude for the process with external momenta  $k_1, k_2 \cdots k_{n-1}$  and  $p$ . If one of the particles with momentum  $p$  splits into two with momenta  $p_1$  and  $p_2$ , then the new

amplitude would be

$$\mathcal{M}_{n+1} = \mathcal{M}_n (-ig) \frac{i}{p^2 + i\epsilon}. \quad (2.2)$$

Since  $p_1$  and  $p_2$  are on-shell massless external momenta, the denominator of the propagator in Eq. (2.2) is

$$p^2 = (p_1 + p_2)^2 = 2p_1 \cdot p_2 = 2E_1 E_2 (1 - \cos(\theta)), \quad (2.3)$$

where  $\theta$  is the spatial angle between  $p_1$  and  $p_2$ . There are two possible cases that would lead to divergences due to the splitting: either  $E_1$  (equivalently for  $E_2$ ) goes to zero or the splitting angle  $\theta$  goes to zero. The divergence that corresponds to the former is called a soft divergence and the latter a collinear divergence.

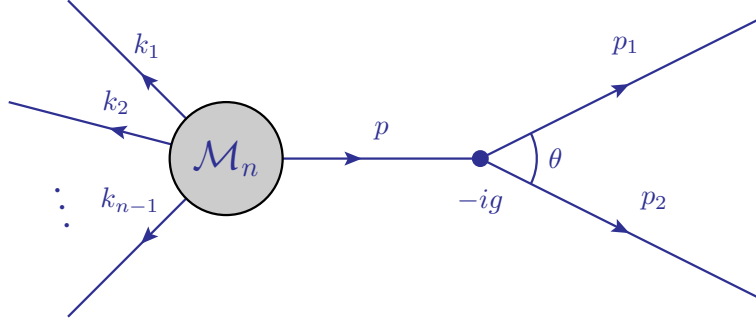


Figure 3: Infrared divergence in  $\phi^3$  theory. One of the external particles with momentum  $p$  splits into two with momenta  $p_1$  and  $p_2$ .

Note that  $\mathcal{M}_{n+1}$  itself is not an observable, so the analysis of the divergences of  $\mathcal{M}_{n+1}$  is more subtle. Let us discuss cross sections associated with the production of  $n$  and  $n + 1$  particles, respectively. With the splitting, the cross section

$$\sigma_n = \int \prod_{i=1}^{n-1} \frac{d^{D-1}k_i}{(2\pi)^{D-1}2E_{k_i}} \int \frac{d^{D-1}p}{(2\pi)^{D-1}2E_p} (2\pi)^D \delta^D \left( p + \sum_i k_i \right) |\mathcal{M}_n|^2 \quad (2.4)$$

changes to

$$\sigma_{n+1} = \int \prod_{i=1}^{n-1} \frac{d^{D-1}k_i}{(2\pi)^{D-1}2E_{k_i}} \int \frac{d^{D-1}p_1}{(2\pi)^{D-1}2E_1} \frac{d^{D-1}p_2}{(2\pi)^{D-1}2E_2} (2\pi)^D \delta^D \left( p_1 + p_2 + \sum_i k_i \right) |\mathcal{M}_{n+1}|^2, \quad (2.5)$$

where

$$|\mathcal{M}_{n+1}| = |\mathcal{M}_n| \frac{g^2}{(p_1 + p_2)^4} = |\mathcal{M}_n| \frac{g^2}{(E_1 E_2 (1 - \cos(\theta)))^2}, \quad (2.6)$$

with the understanding that  $p$  is not on-shell as it is in Eq. (2.4). There is one trick that facilitates the comparison between  $\sigma_n$  and  $\sigma_{n+1}$ . Inserting the identity

$$1 = \int \frac{d^{D-1}p}{(2\pi)^{D-1}2E_p} (2\pi)^{D-1} \delta^{D-1}(p_1 + p_2 - p) \quad (2.7)$$

in Eq. (2.5), one obtains

$$\begin{aligned} \sigma_{n+1} = & g^2 \int \frac{d^{D-1}p}{(2\pi)^{D-1}2E_p} \int \frac{d^{D-1}p_1}{(2\pi)^{D-1}2E_1} \frac{d^{D-1}p_2}{(2\pi)^{D-1}2E_2} \left( \int \prod_{i=1}^{n-1} \frac{d^{D-1}k_i}{(2\pi)^{D-1}2E_{k_i}} \right. \\ & \left. \times (2\pi)^D \delta^D(p_1 + p_2 + \sum_i^{n-1} k_i) |\mathcal{M}_n|^2 \right) (2\pi)^{D-1} \delta^{D-1}(p - p_1 - p_2) \frac{1}{(E_1 E_2 (1 - \cos(\theta)))^2}. \end{aligned} \quad (2.8)$$

To study the phase space integration with respect to  $p_1$  and  $p_2$ , we first integrate out the angular coordinates except for  $\theta$  which results in

$$\int d^{D-1}p_1 d^{D-1}p_2 \cdots = \Omega_{D-3} \int dE_1 dE_2 E_1^{D-2} E_2^{D-2} d\theta \sin^{D-3}(\theta) \cdots, \quad (2.9)$$

where  $\Omega_{D-3}$  is the volume of a unit  $(D-3)$ -sphere. In the soft-collinear limit ( $\theta \rightarrow 0$ ,  $E_1$  or  $E_2 \rightarrow 0$ ),  $d^{D-1}p_1$  and  $d^{D-1}p_2$  with  $d\Omega_{D-3}$  integrated out give rise to

$$\int dE_1 dE_2 d\theta E_1^{D-5} E_2^{D-5} \theta^{D-7}. \quad (2.10)$$

Eq. (2.10) shows that the infrared behavior of the splitting depends on the space-time dimension. In particular, if  $D > 6$  there is no infrared divergence. If  $4 < D \leq 6$ , the infrared divergence comes from the angular integration. If  $D \leq 4$ , both the energy and angular integration give rise to infrared divergences.

The origin of the infrared divergences considered above is the degeneracy of external states. Experimentally, one can not distinguish between a massless particle and the collection of collinear massless particles with the same energy, nor can one distinguish between a massless particle and a massless particle accompanied by a collection of very soft massless particles. This means that neither  $\sigma_n$  nor  $\sigma_{n+1}$  is an observable. One should combine  $\sigma_n$  with

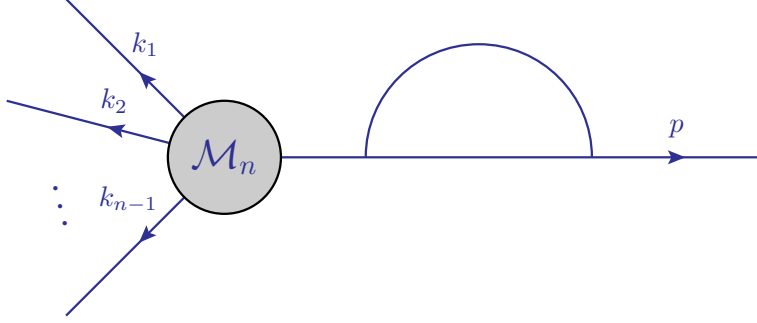


Figure 4: Virtual correction to  $\mathcal{M}_n$  needed to cancel infrared divergences that appear in  $\mathcal{M}_{n+1}$ .

part of  $\sigma_{n+1}$  where the phase space integration in Eq. (2.8) are constrained to be in the soft or collinear regions for  $p_1$  and  $p_2$ . In addition, one should also includes virtual corrections to  $\mathcal{M}_n$  by considering diagrams such as that in Figure 4.

It is interesting to mention that the infrared divergence of the scattering amplitude discussed above is due to choosing initial and final states as single particle states. It is possible to define infrared finite amplitudes by carefully defining asymptotic states first [22, 23, 24]. For instance, in QED, the infrared divergence comes from the emission of an indefinite number of soft-photons. If one defines initial and final states as the superposition of a electron state and an infinite number of photon states (coherent states), an infrared finite scattering amplitude can be defined [22]. In QCD, the degeneracy from color degree of freedom and additional gluon vertices make the infrared divergences much more dire than that in QED.

### 2.1.2 SCET for $\phi^3$ theories

In this section, we review SCET for  $\phi^3$  theory and its application to the factorization of the Sudakov form factor that are discussed in more details in Ref. [25]. The SCET for massless  $\phi^3$  theory is much easier to deal with since we don't need to worry about the non-abelian gauge fields associated with  $SU(3)$  in QCD. Nevertheless it incorporates the main

ingredients of SCET in general. This will also help us set up notations and nomenclatures that are commonly used in SCET in the context of QCD.

**2.1.2.1 SCET Lagrangian for  $\phi^3$  theory** The discussion in Section 2.1.1 on infrared divergences show that soft and collinear radiations of  $\phi$  particles are enhanced. So there could be a jet like cascade if the  $\phi$  particle propagates in some material with which it interacts. The dynamics of the jet (i.e., the interaction of particles inside the jet or interactions with other jets) could be described by an effective field whose momentum fluctuations are small around the jet direction.

In SCET, one usually works with light-cone coordinates. Let  $\hat{n}$  be a unit 3-vector (it could be a jet direction), and let  $n = (1, \hat{n})$  and  $\bar{n} = (1, -\hat{n})$ . Then  $n^2 = \bar{n}^2 = 0$ , i.e.,  $n$  and  $\bar{n}$  are two light-cone vectors with opposite directions. Any 4-vector  $k$  can be decomposed along  $n$ ,  $\bar{n}$ , and a direction perpendicular to both  $n$  and  $\bar{n}$  as follows:

$$k^\mu = n \cdot k \frac{\bar{n}^\mu}{2} + \bar{n} \cdot k \frac{n^\mu}{2} + k_\perp^\mu. \quad (2.11)$$

The light-cone components of  $k$  are denoted as

$$(k^+, k^-, k_\perp) \equiv (n \cdot k, \bar{n} \cdot k, k_\perp). \quad (2.12)$$

Let  $q$  be another vector, then  $k^2$  and the product of  $k$  and  $q$  have the following expressions:

$$\begin{aligned} k^2 &= k^+ k^- + k_\perp^2 \\ k \cdot q &= \frac{1}{2} k^+ q^- + \frac{1}{2} k^- q^+ + k_\perp \cdot p_\perp. \end{aligned} \quad (2.13)$$

where

$$\begin{aligned} k_\perp^2 &= -\vec{k}_\perp^2 \\ k_\perp \cdot p_\perp &= -\vec{k}_\perp \cdot \vec{p}_\perp \end{aligned} \quad (2.14)$$

If  $p$  is a momentum collinear to  $n$  and  $l$  collinear to  $\bar{n}$ , then  $p$  and  $l$  scale as

$$\begin{aligned} (p^+, p^-, p_\perp) &\sim (\lambda^2, 1, \lambda) \\ (l^+, l^-, l_\perp) &\sim (1, \lambda^2, \lambda), \end{aligned} \quad (2.15)$$

where  $\lambda \ll 1$  is the power counting parameter for SCET with  $\lambda \sim p_\perp/p^- \sim l_\perp/l^+$ . The scaling for  $p$  and  $l$  in Eq. (2.15) can be obtained as follows. Let  $p$  be on-shell as it should be if it's a final state particle in a jet, i.e.,  $p^2 = 0$ . Since the direction of  $\vec{p}$  is close to  $\hat{n}$ ,

$$\begin{aligned} p^- &= \bar{n} \cdot p = p^0 + \hat{n} \cdot \vec{p}, \\ p^+ &= n \cdot p = p^0 - \hat{n} \cdot \vec{p}, \end{aligned} \tag{2.16}$$

and

$$p^- \sim 2p^0 \gg p^+. \tag{2.17}$$

In addition  $p^- \gg p_\perp$ , otherwise  $p$  is by no means collinear to  $\hat{n}$ . The on-shell condition requires

$$p^2 = 0 \quad \Rightarrow \quad p^+ = \frac{\vec{p}_\perp^2}{p^-}. \tag{2.18}$$

Let  $\lambda \sim p_\perp/p^-$ . Then

$$(p^+, p^-, p_\perp) = p^- \left( \frac{p^+}{p^-}, 1, \frac{p_\perp}{p^-} \right) \sim p^- (\lambda^2, 1, \lambda) \tag{2.19}$$

which is the scaling shown in Eq. (2.15). By exactly the same argument, one can obtain the scaling for  $l$  in Eq. (2.15).

Assume we are interested in a back to back di-jet process, with one along  $n$  and the other  $\bar{n}$ . Then there are two collinear fields we need to construct. Let  $\phi_c$  and  $\phi_{\bar{c}}$  be fields (or modes of  $\phi$ ) collinear to  $n$  and  $\bar{n}$  respectively. In addition, one needs to introduce another mode  $\phi_s$  to mediate the interaction between  $\phi_c$  and  $\phi_{\bar{c}}$ .  $\phi_s$  should be soft. More specifically, it should not take  $\phi_c$  or  $\phi_{\bar{c}}$  off the scaling shown in Eq. (2.15). The momentum of  $\phi_s$  scales as

$$(p_s^+, p_s^-, p_{s\perp}) \sim (\lambda^2, \lambda^2, \lambda^2). \tag{2.20}$$

When interacting with collinear modes, this soft momentum scaling leaves the collinear scalings in Eq. (2.15) unchanged. One natural question to ask is whether other soft modes could be important. This is dependent on what processes we are interested in. For the presently considered di-jet process, this is the only soft mode that is relevant. For instance, modes with momentum scaling as  $(\lambda^3, \lambda^3, \lambda^3)$  have already been included in the soft modes



whose momentum scales as  $(\lambda^2, \lambda^2, \lambda^2)$ . Or if one insists on including those modes in the calculation, the Feynman diagrams involving those modes give zero contributions.

The effective Lagrangian has the following form including the interesting modes in the di-jet process:

$$\begin{aligned}\mathcal{L} = & \frac{1}{2}\partial_\mu\phi_c\partial^\mu\phi_c - \frac{g}{3!}\phi_c^3 + \dots \\ & + \frac{1}{2}\partial_\mu\phi_{\bar{c}}\partial^\mu\phi_{\bar{c}} - \frac{g}{3!}\phi_{\bar{c}}^3 + \dots \\ & + \frac{1}{2}\partial_\mu\phi_s\partial^\mu\phi_s - \frac{g}{3!}\phi_s^3 + \dots \\ & + \mathcal{L}_{c,\bar{c},s},\end{aligned}\tag{2.21}$$

where the  $\dots$ 's are higher order power corrections when matching to the full  $\phi^3$  theory (power counting of fields is discussed below). The leading order interactions for  $\mathcal{L}_{c,\bar{c},s}$  are

$$\mathcal{L}_{c,\bar{c},s} = -\frac{g}{2}\phi_c^2\phi_s - \frac{g}{2}\phi_{\bar{c}}^2\phi_s + \dots.\tag{2.22}$$

There are no terms, for example, like  $[\phi_c\phi_s^2]$  in  $\mathcal{L}_{c,\bar{c},s}$  at leading power simply because of momentum conservation: a collinear particle can not decay into two soft particles, otherwise it violates the power counting rules in Eq. (2.15) and Eq. (2.20).

Now consider the power counting of fields, which is the first thing one needs to investigate when constructing EFT operators. Let the space-time dimension be 6 (so that the  $\phi^3$  theory is renormalizable). For collinear fields,

$$\int d^6p \sim \lambda^6, \quad p^2 \sim \lambda^2,\tag{2.23}$$

where  $dp^- \sim 1$ ,  $dp^+ \sim \lambda^2$ , and  $d^4p_\perp \sim \lambda^4$ . As a result,

$$\langle 0|T\phi_c(x)\phi_c(0)|0\rangle \sim \int d^6p e^{-ipx} \frac{i}{p^2} \sim \lambda^4 \quad \Rightarrow \quad \phi_c \sim \lambda^2.\tag{2.24}$$

Similarly, for soft fields,

$$\int d^6p \sim \lambda^{12}, \quad p^2 \sim \lambda^4,\tag{2.25}$$

where  $dp^- \sim \lambda^2$ ,  $dp^+ \sim \lambda^2$ , and  $d^4p_\perp \sim \lambda^8$ . As a result,

$$\langle 0|T\phi_s(x)\phi_s(0)|0\rangle \sim \int d^6p e^{-ipx} \frac{i}{p^2} \sim \lambda^8 \quad \Rightarrow \quad \phi_s \sim \lambda^4.\tag{2.26}$$

For the operators in the Lagrangian in Eq. (2.21)

$$\begin{aligned}
\int d^6x (\partial_\mu \phi_c)^2 &\sim \frac{1}{\lambda^6} (\lambda^2) (\lambda^4) = \lambda^0, \\
\int d^6x (\partial_\mu \phi_{\bar{c}})^2 &\sim \frac{1}{\lambda^6} (\lambda^2) (\lambda^4) = \lambda^0, \\
\int d^6x (\partial_\mu \phi_s)^2 &\sim \frac{1}{\lambda^{12}} (\lambda^4) (\lambda^8) = \lambda^0, \\
\int d^6x \phi_c^3 &\sim \frac{1}{\lambda^6} (\lambda^2)^3 = \lambda^0, \\
\int d^6x \phi_s^3 &\sim \frac{1}{\lambda^{12}} (\lambda^4)^3 = \lambda^0,
\end{aligned} \tag{2.27}$$

where we used  $x^\mu \sim 1/p_\mu$  according to the uncertainty principle and  $\partial_\mu \sim p_\mu$ . Eq. (2.27) shows that the terms explicitly written out in Eq. (2.21) are the terms at leading power. For the interaction term  $\mathcal{L}_{c,\bar{c},s}$  in Eq. (2.22),

$$\begin{aligned}
\int d^6x \phi_c^2 \phi_s &\sim \frac{1}{\lambda^6} (\lambda^4) (\lambda^4) = \lambda^2, \\
\int d^6x \phi_{\bar{c}}^2 \phi_s &\sim \frac{1}{\lambda^6} (\lambda^4) (\lambda^4) = \lambda^2,
\end{aligned} \tag{2.28}$$

where the scaling  $\int d^6x$  is determined by  $\phi_c$  (or  $\phi_{\bar{c}}$ ) since the summation of collinear and soft momenta has a collinear scaling. So at leading power  $\mathcal{O}(\lambda^0)$  there is no interaction between the two collinear sectors, and up to  $\mathcal{O}(\lambda^2)$   $\mathcal{L}_{c,\bar{c},s}$  should be included.

**2.1.2.2 Factorization of the Sudakov Factor** Now let's consider the factorization of Sudakov form factor in the context of  $\phi^3$  theory. The form factor refers to the matrix element

$$G = \langle l, p | j(0) | 0 \rangle, \tag{2.29}$$

where  $j(x)$  is a current vector defined as

$$j(x) \equiv \phi^2(x), \tag{2.30}$$

and  $l$  and  $p$  are external momenta. Let  $l$  (associated with  $\phi_c$ ) and  $p$  (associated with  $\phi_{\bar{c}}$ ) be collinear momenta along  $n$  and  $\bar{n}$ , respectively. This form factor is relevant, for instance, to the di-jet process. Diagrammatically the form factor involves studying the diagram shown in Figure 5.

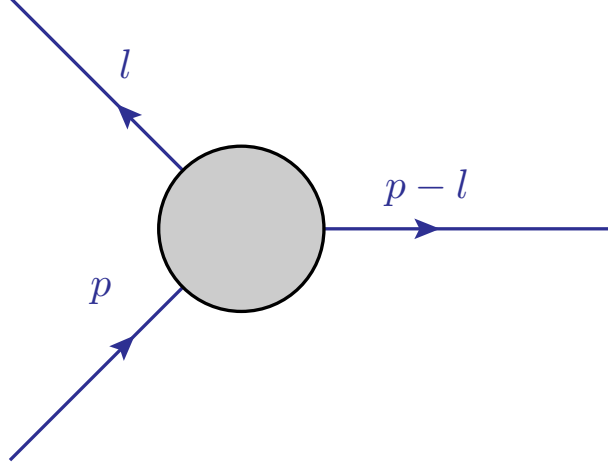


Figure 5: Sudakov form factor for  $\phi^3$  theory.

Take the power counting parameter to be

$$\lambda^2 \sim P^2/Q^2 \sim L^2/Q^2, \quad (2.31)$$

with  $\lambda \ll 1$ ,  $P^2 \equiv -p^2$ ,  $L^2 \equiv -l^2$ , and  $Q^2 \equiv -(l-p)^2$ . Note that this is consistent with the momentum scaling discussed in Section 2.1.2.1, since

$$\begin{aligned} p_\perp^2/p^{-2} &\sim p^2/p^{-2} \sim \lambda^2, \\ p^{-2}/p^{-2} &\sim Q^2/p^{-2} \sim 1. \end{aligned} \quad (2.32)$$

To study the Sudakov Form Factor, we need to match the full theory current to the effective theory current. First, note that derivatives of collinear fields along their light-cone directions are not power suppressed:

$$\phi_c(x + s\bar{n}) = \sum \frac{s^i}{i!} (\bar{n} \cdot \partial)^i \phi \quad \text{with} \quad \bar{n} \cdot \partial \rightarrow \bar{n} \cdot p \sim \lambda^0. \quad (2.33)$$

So at leading power, we should do the following matching:

$$j(x) = [\phi(x)]^2 \rightarrow j_{\text{SCET}}(x) = \int ds \int dt C(s, t, \mu) \phi_c(x + s\bar{n}) \phi_{\bar{c}}(x + tn) + \mathcal{O}(\lambda), \quad (2.34)$$

where  $C(s, t, \mu)$  is the Wilson coefficients associated with the SCET composite operator  $(\phi_c \phi_{\bar{c}})$  that can be calculated perturbatively by matching SCET diagrams with diagram in the full  $\phi^3$  theory. The power counting of  $[\phi_c(x) \phi_{\bar{c}}(x)]$  is

$$\int d^6 x \phi_c \phi_{\bar{c}} \sim \frac{1}{\lambda^4} (\lambda^2) (\lambda^2) = \lambda^0 \quad (2.35)$$

where  $\int d^6 x \sim \lambda^{-4}$  since the combination of momenta for  $\phi_c$  and  $\phi_{\bar{c}}$  gives

$$(p_c^+ + p_{\bar{c}}^+, p_c^- + p_{\bar{c}}^-, p_{c\perp} + p_{\bar{c}\perp}) \sim (1, 1, \lambda), \quad (2.36)$$

and so  $\int d^6 p \sim \lambda^4$  and as a result  $\int d^6 x \sim \lambda^{-4}$ . The power counting of  $[\phi_c \phi_{\bar{c}}]$  justifies the notation  $\mathcal{O}(\lambda)$  in Eq. (2.34).

At tree level (leading order in perturbation) and up to  $\mathcal{O}(\lambda)$  (leading order in power correction), the SCET current operator is

$$j_{\text{SCET}}(x) = \phi_c(x) \phi_{\bar{c}}(x) + \mathcal{O}(\lambda, g^2), \quad (2.37)$$

i.e.,

$$C(s, t, \mu) = \delta(s) \delta(t) + \mathcal{O}(\lambda, g^2). \quad (2.38)$$

With  $j_{\text{SCET}}$  in hand, one can factorize the Sudakov form factor as follows:

$$\begin{aligned} G(p, l, \mu) &= \int dx_1 dx_2 e^{-ipx_1 + ilx_2} \langle 0 | T[\phi(x_1) j_{\text{SCET}}(0) \phi(x_2)] | 0 \rangle \\ &= \int dx_1 dx_2 e^{-ipx_1 + ilx_2} \langle 0 | T[\phi_c(x_1) j_{\text{SCET}}(0) \phi_{\bar{c}}(x_2)] | 0 \rangle \\ &= \int dx_1 dx_2 e^{-ipx_1 + ilx_2} \int ds dt C(s, t, \mu) \langle 0 | T[\phi_c(x_1) \phi_c(s\bar{n})] | 0 \rangle \langle 0 | T[\phi_{\bar{c}}(t\bar{n}) \phi_{\bar{c}}(x_2)] | 0 \rangle \\ &= \tilde{\mathcal{C}}(\bar{n} \cdot p, n \cdot l, \mu) \mathcal{J}(p^2, \mu) \mathcal{J}(l^2, \mu) \end{aligned} \quad (2.39)$$

From the second line to the third line, the fact was used that in the effective theory Lagrangian, different collinear sectors are decoupled at leading power. In the last line,  $\mathcal{J}(p^2, \mu) \equiv \int d^d x_1 e^{-ipx_1} \langle 0 | \phi_c(x_1) \phi_c(0) | 0 \rangle$  and  $\mathcal{J}(l^2, \mu) \equiv \int d^d x_1 e^{-ilx_2} \langle 0 | \phi_{\bar{c}}(x_2) \phi_{\bar{c}}(0) | 0 \rangle$  are called Jet Functions, and  $\tilde{\mathcal{C}}(\bar{n} \cdot p, n \cdot l, \mu) \equiv \int ds \int dt C(s, t, \mu) e^{-isp \cdot \bar{n} + itl \cdot n}$  is called the Hard Function.

### 2.1.3 SCET Lagrangian for QCD

Just like  $\phi^3$  theory, for each collinear sector, collinear fields are introduced for both quarks and gluons. However, the power counting rules are much more complicated with different components of quark spinors and gauge field vectors scaling differently. In addition the  $SU(3)$  gauge symmetry is realized in a complicated fashion.

First, consider collinear quark field  $\psi_n$  in the  $n$  direction.<sup>1</sup> To figure out the power counting rules for different components of  $\psi_n$ , consider the Dirac equation in momentum space for a massless fermion (such as a collinear particle in high energy jets):

$$\not{p}\psi_n = 0 = \left( \frac{\bar{n} \cdot p}{2} \not{n} + \frac{n \cdot p}{2} \not{\bar{n}} + \not{p}_\perp \right) \psi_n. \quad (2.40)$$

Since  $\psi_n$  is supposed to be collinear to  $n$  and  $n^\mu \sim 1$ ,

$$\frac{\bar{n} \cdot p}{2} \not{n} \sim p^0, \quad \frac{n \cdot p}{2} \not{\bar{n}} \sim p^0 \lambda^2, \quad \not{p}_\perp \sim p^0 \lambda. \quad (2.41)$$

As a result, Eq. (2.40) reduces to

$$\not{p}\psi_n = 0 \quad \text{at leading power.} \quad (2.42)$$

This motivates the decomposition of  $\psi_n$  into two parts

$$\psi_n = \xi_n + \eta_{\bar{n}}, \quad (2.43)$$

where

$$\xi_n \equiv \mathcal{P}_n \psi_n, \quad \eta_{\bar{n}} \equiv \mathcal{P}_{\bar{n}} \psi_n, \quad (2.44)$$

and

$$\mathcal{P}_n \equiv \frac{\not{n} \not{\bar{n}}}{4}, \quad \mathcal{P}_{\bar{n}} \equiv \frac{\not{\bar{n}} \not{n}}{4}. \quad (2.45)$$

$\mathcal{P}_n$  and  $\mathcal{P}_{\bar{n}}$  are projection operators that satisfy

$$\mathcal{P}_n^2 = \mathcal{P}_n, \quad \mathcal{P}_{\bar{n}}^2 = \mathcal{P}_{\bar{n}}, \quad \mathcal{P}_n \mathcal{P}_{\bar{n}} = 0, \quad \mathcal{P}_n + \mathcal{P}_{\bar{n}} = \mathbb{1}. \quad (2.46)$$

One can immediately see that

$$\not{n} \xi_n = 0, \quad \not{\bar{n}} \eta_{\bar{n}} = 0. \quad (2.47)$$

---

<sup>1</sup>Conventions in light-cone coordinates are setup in Section 2.1.2.1.

To see what  $\mathcal{P}_n$  and  $\mathcal{P}_{\bar{n}}$  do when they operate on quark spinors, let  $n = (1, 0, 0, 1)$  and work in the Dirac representation. In the Dirac representation, the  $\gamma$  matrices are

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}. \quad (2.48)$$

So

$$\not{n} = \gamma^0 - \gamma^3 = \begin{pmatrix} \mathbb{1} & -\sigma^3 \\ \sigma^3 & -\mathbb{1} \end{pmatrix}, \quad \not{\bar{n}} = \gamma^0 + \gamma^3 = \begin{pmatrix} \mathbb{1} & \sigma^3 \\ -\sigma^3 & -\mathbb{1} \end{pmatrix}, \quad (2.49)$$

and

$$\mathcal{P}_n = \frac{\not{n}\not{\bar{n}}}{4} = \frac{1}{2} \begin{pmatrix} \mathbb{1} & \sigma^3 \\ \sigma^3 & \mathbb{1} \end{pmatrix}, \quad \mathcal{P}_{\bar{n}} = \frac{\not{\bar{n}}\not{n}}{4} = \frac{1}{2} \begin{pmatrix} \mathbb{1} & -\sigma^3 \\ -\sigma^3 & \mathbb{1} \end{pmatrix}. \quad (2.50)$$

The quark and anti-quark spinors are

$$u_n \propto \begin{pmatrix} \mathcal{U} \\ \sigma^3 \mathcal{U} \end{pmatrix}, \quad v_n \propto \begin{pmatrix} \sigma^3 \mathcal{V} \\ \mathcal{V} \end{pmatrix}, \quad (2.51)$$

where  $\mathcal{U}$  and  $\mathcal{V}$  are two-component vectors corresponding to different spin states for particle and anti-particles, respectively.

From Eq. (2.50) and (2.51),

$$\mathcal{P}_n u_n = u_n, \quad \mathcal{P}_n v_n = v_n, \quad \mathcal{P}_{\bar{n}} u_n = \mathcal{P}_{\bar{n}} v_n = 0. \quad (2.52)$$

So the projection operator  $\mathcal{P}_n$  preserves both the particle and the anti-particle components of the collinear fields. This is different than the NRQCD Lagrangian to be discussed in Section 2.2.2.

Now consider the power counting of  $\xi_n(x)$  and  $\eta_n(x)$  by looking at their propagators. The following identities are useful for evaluating the power countings of  $\xi_n$  and  $\eta_{\bar{n}}$ :

$$\begin{aligned} \not{n} \not{p}_\perp &= -\not{p}_\perp \not{n}, & \not{\bar{n}} \not{p}_\perp &= -\not{p}_\perp \not{\bar{n}}, \\ \not{n} \not{n} &= \not{\bar{n}} \not{\bar{n}} = 0, & \{\not{n}, \not{\bar{n}}\} &= 4, \\ \mathcal{P}_n \not{n} &= \not{n}, & \mathcal{P}_{\bar{n}} \not{\bar{n}} &= \not{\bar{n}}, & \not{n} \mathcal{P}_n &= \not{n}, & \not{\bar{n}} \mathcal{P}_{\bar{n}} &= \not{\bar{n}}, \\ \mathcal{P}_n \not{p} \mathcal{P}_{\bar{n}} &= \frac{\bar{n} \cdot p}{2} \not{n}, & \mathcal{P}_{\bar{n}} \not{p} \mathcal{P}_n &= \frac{n \cdot p}{2} \not{\bar{n}}, \end{aligned} \quad (2.53)$$

For  $\xi_n(x)$ ,

$$\begin{aligned}
\langle 0 | \xi_n(x) \bar{\xi}_n(0) | 0 \rangle &= \langle 0 | \mathcal{P}_n \psi_n(x) \bar{\psi}_n(0) \mathcal{P}_{\bar{n}} | 0 \rangle \\
&= \mathcal{P}_n \langle 0 | \psi_n(x) \bar{\psi}_n(0) | 0 \rangle \mathcal{P}_{\bar{n}} \\
&= \mathcal{P}_n \left[ \int \frac{d^4 p}{(2\pi)^4} \frac{i \not{p}}{p^2} e^{-ip \cdot x} \right] \mathcal{P}_{\bar{n}} \\
&= \int \frac{d^4 p}{(2\pi)^4} \frac{1}{2} \frac{i \bar{n} \cdot p \not{p}}{p^2} e^{-ip \cdot x} \\
&\sim (\lambda^4) (\lambda^0) \frac{1}{\lambda^2} \\
&= \lambda^2,
\end{aligned} \tag{2.54}$$

where

$$\mathcal{P}_n^\dagger = \mathcal{P}_{\bar{n}}, \tag{2.55}$$

and identities in Eq. (2.53) have been used. This says that the power counting of  $\xi_n$  is

$$\xi_n(x) \sim \lambda. \tag{2.56}$$

We can do similar manipulations for  $\eta_{\bar{n}}$ :

$$\begin{aligned}
\langle 0 | \eta_{\bar{n}}(x) \bar{\eta}_{\bar{n}}(0) | 0 \rangle &= \langle 0 | \mathcal{P}_{\bar{n}} \psi_n(x) \bar{\psi}_n(0) \mathcal{P}_n | 0 \rangle \\
&= \mathcal{P}_{\bar{n}} \langle 0 | \psi_n(x) \bar{\psi}_n(0) | 0 \rangle \mathcal{P}_n \\
&= \mathcal{P}_{\bar{n}} \left[ \int \frac{d^4 p}{(2\pi)^4} \frac{i \not{p}}{p^2} e^{-ip \cdot x} \right] \mathcal{P}_n \\
&= \int \frac{d^4 p}{(2\pi)^4} \frac{1}{2} \frac{i n \cdot p \not{p}}{p^2} e^{-ip \cdot x} \\
&\sim (\lambda^4) (\lambda^2) \frac{1}{\lambda^2} \\
&= \lambda^4,
\end{aligned} \tag{2.57}$$

which shows that the scaling of  $\eta_{\bar{n}}$  is

$$\eta_{\bar{n}}(x) \sim \lambda^2. \tag{2.58}$$

By similar arguments as those in Eq. (2.54) and (2.57), or by noticing that gauge fields are always associated with space-time derivatives  $\partial_\mu$ , one obtains the scalings of  $n$ -collinear gluon fields:

$$(n \cdot A_n, \bar{n} \cdot A_n, A_{n\perp}) \sim (\lambda^2, 1, \lambda). \quad (2.59)$$

To find the SCET Lagrangian for the  $n$ -collinear field, plug the decomposition of  $\psi_n$  in Eq. (2.43) into the full QCD Lagrangian (massless quark):

$$\begin{aligned} \mathcal{L}_n &= \bar{\psi}_n i \not{D} \psi_n \\ &= (\bar{\xi}_n + \bar{\eta}_{\bar{n}}) \left( \frac{i \bar{n} \cdot D}{2} \not{n} + \frac{i n \cdot D}{2} \not{\bar{n}} + i \not{D}_\perp \right) (\xi_n + \eta_{\bar{n}}) \\ &= \bar{\xi}_n \frac{\not{n}}{2} i n \cdot D \xi_n + \bar{\eta}_{\bar{n}} \frac{\not{\bar{n}}}{2} i \bar{n} \cdot D \eta_{\bar{n}} + \bar{\xi}_n i \not{D}_\perp \eta_{\bar{n}} + \bar{\eta}_{\bar{n}} i \not{D}_\perp \xi_n. \end{aligned} \quad (2.60)$$

Since  $\eta_{\bar{n}}(x)$  is power suppressed compared to  $\xi_n(x)$  (shown in Eq. (2.56) and (2.58)), we can integrate out  $\eta_{\bar{n}}$  at leading power using the equation of motion. From the Lagrangian Eq. (2.60), the equation of motion for  $\eta_{\bar{n}}$  is

$$\frac{\not{n}}{2} i \bar{n} \cdot D \eta_{\bar{n}} + i \not{D}_\perp \xi_n = 0. \quad (2.61)$$

After some algebra (using identities in Eq. (2.53)),

$$\eta_{\bar{n}} = \frac{1}{i \bar{n} \cdot D} i \not{D}_\perp \bar{n} \xi_n. \quad (2.62)$$

Plugging this back into Eq. (2.60), one obtains

$$\boxed{\mathcal{L}_n = \bar{\xi}_n \left( i n \cdot D + i \not{D}_\perp \frac{1}{i \bar{n} \cdot D} i \not{D}_\perp \right) \frac{\not{n}}{2} \xi_n.} \quad (2.63)$$

This is the Lagrangian for collinear quark (jet) field.  $n$  is an arbitrary jet direction. For each jet direction, one collinear quark field can be introduced. The interactions with gluon fields are determined by the covariant derivative

$$D_\mu = \partial_\mu - i g_s T^a A_\mu^a \quad (2.64)$$

which will be discussed in more details in the next section.



### 2.1.4 Gauge Transformations

The structure of gauge interaction is determined by the gauge invariance associated with a certain gauge group ( $SU(3)$  for QCD). In Eq. (2.63), the gauge field  $A_\mu$  in the covariant derivative  $D_\mu$  should transform accordingly so as to match the local gauge transformation of the collinear quark field  $\xi_n$ . Let

$$U(x) = \exp[i\alpha^a(x)T^a] \quad (2.65)$$

be a  $SU(3)$  gauge transformation on  $\xi_n$ . Unlike full QCD, the choice of  $\alpha^a(x)$  is not arbitrary in that the transformation should not upset the power counting of  $\xi_n$ . For instance, if  $\partial_\mu \alpha(x)$  scales like a hard momentum

$$(n \cdot \partial \alpha, \bar{n} \cdot \partial \alpha, \not{\partial}_\perp \alpha) \sim p^0(1, 1, 1), \quad (2.66)$$

where  $p^0$  is the energy scale of  $\xi_n$ , then the gauge transformation will carry  $\xi_n$  to hard scaling which is supposed to have been integrated out already in SCET.

Depending on what physical processes is under consideration, the interesting gauge field modes may be quite different (See Section 3.2 for a working example of a different gauge field mode not considered here). In this section, we focus on two modes that correspond to collinear and ultra-soft gauge transformation, respectively. The scalings of collinear and ultra-soft gauge transformations are

$$(n \cdot \partial \alpha_n, \bar{n} \cdot \partial \alpha_n, \not{\partial}_\perp \alpha_n) \sim p^0(\lambda^2, 1, \lambda), \quad (2.67)$$

$$(n \cdot \partial \alpha_s, \bar{n} \cdot \partial \alpha_s, \not{\partial}_\perp \alpha_s) \sim p^0(\lambda^2, \lambda^2, \lambda^2), \quad (2.68)$$

respectively. The QCD effective theory where modes with collinear and ultra-soft scaling are relevant is called SCET-I. The term ‘ultra-soft’ scaling was introduced to distinguish it from the soft scaling

$$(n \cdot p_{\text{soft}}, \bar{n} \cdot p_{\text{soft}}, p_{\text{soft}\perp}) \sim (\lambda, \lambda, \lambda). \quad (2.69)$$

The QCD effective theory where modes with collinear and soft scaling are relevant is called SCET-II. A typical example is  $B \rightarrow D\pi$  decay process for which the power counting parameter is  $\lambda \sim \Lambda_{QCD}/m_B$ .

Let us first consider the ultra-soft gauge transformation

$$\xi_n(x) \rightarrow U_s(x) \xi_n(x) = \exp [i\alpha_s^a(x) T^a] \xi_n(x). \quad (2.70)$$

Gauge invariance of the Lagrangian requires an ultra-soft gauge field  $A_s$  in the covariant derivative

$$D_s^\mu = \partial^\mu - ig_s A_s^{a\mu} T^a \quad (2.71)$$

with  $A_s$  transforming as

$$A_s^\mu \rightarrow U_s A_s^\mu U_s^\dagger + \frac{i}{g_s} U_s [\partial^\mu, U_s^\dagger], \quad (2.72)$$

where  $A_s^\mu$  is short for  $T^a A_s^{a\mu}$ . The kinetic term for  $A_s$  is the same as that in the full theory

$$-\frac{1}{4} F_{s\mu\nu}^a F_s^{a\mu\nu}, \quad (2.73)$$

where

$$F_{s\mu\nu} = F_{s\mu\nu}^a T^a \equiv \frac{i}{g} [D_{s\mu}, D_{s\nu}]. \quad (2.74)$$

Similarly, to preserve the following collinear gauge transformation invariance in the Lagrangian,

$$\xi_n(x) \rightarrow U_n(x) \xi_n(x) = \exp [i\alpha_n^a(x) T^a] \xi_n(x), \quad (2.75)$$

there should be a collinear mode  $A_n^\mu$  in the gauge sector that transforms as

$$A_n^\mu \rightarrow U_n A_n^\mu U_n^\dagger + \frac{i}{g_s} U_n [\mathcal{D}_n^\mu, U_n^\dagger], \quad (2.76)$$

where

$$\mathcal{D}_n^\mu = \partial^\mu - ig_s \frac{\bar{n}^\mu}{2} n \cdot A_s. \quad (2.77)$$

Note that the comutator in Eq. (2.76) is  $[\mathcal{D}_n^\mu, U_n^\dagger]$  instead of  $[\partial^\mu, U_n^\dagger]$  such as that in Eq. (2.72). This is due to the scaling behavior of the  $\bar{n}$  component of collinear and ultra-soft gluons

$$n \cdot A_n \sim \lambda^2, \quad n \cdot A_s \sim \lambda^2. \quad (2.78)$$

It seems that one needs to transform one of the ultra-soft components  $n \cdot A_s$  under  $U_n$  transformation. One way to deal with this is to insist that  $A_s \rightarrow A_s$  under  $U_n$ . Including both collinear and ultra-soft modes of gluons, the covariant derivative reads

$$D^\mu = \partial^\mu - ig_s A_n^\mu - ig_s A_s^\mu. \quad (2.79)$$

For the  $n$  and  $\perp$  component, there is no problem since  $[\mathcal{D}^\mu, U_n^\dagger]$  and  $[\partial^\mu, U_n^\dagger]$  are the same for those components. For  $\bar{n}$  (small) component, the collinear gauge tranformation invariance requires

$$n \cdot D \rightarrow n \cdot D' = U_n [n \cdot D] U_n^\dagger. \quad (2.80)$$

Since we have chosen  $A_s$  not to be transformed, this leads to

$$n \cdot \partial - ig_s n \cdot A'_n - ig_s n \cdot A_s = U_n [n \cdot \partial - ig_s n \cdot A_n - ig_s n \cdot A_s] U_n^\dagger. \quad (2.81)$$

As a result, the transformation of  $A_n$  is

$$n \cdot A \rightarrow n \cdot A' = U_n n \cdot A_n U_n^\dagger + \frac{i}{g_s} U_n [n \cdot \partial - ig_s n \cdot A_s, U_n^\dagger], \quad (2.82)$$

and the transformation in Eq. (2.76) is justified.

The covariant derivative in Eq. (2.79) involves both collinear and ultra-soft gluon fields. In the light-cone coordinates, only the larger component for  $A_n$  field needs to be kept:

$$\begin{aligned} n \cdot D &= n \cdot \partial - ig_s n \cdot A_n - ig_s n \cdot A_s, \\ \bar{n} \cdot D &= \bar{n} \cdot \partial - ig_s \bar{n} \cdot A_n, \\ \not{D}_\perp &= \not{\partial}_\perp - ig_s \not{A}_{n\perp}. \end{aligned} \quad (2.83)$$

### 2.1.5 Wilson Lines

Wilson lines are ubiquitous in SCET. They are indispensable components in constructing gauge invariant operators and facilitating factorization theorems.

Wilson lines can be introduced in different ways. Mathematically, a Wilson line corresponds to a connection defined in a vector bundle (in the context of QFT).<sup>2</sup> Just like the metric connection in General Relativity, Wilson lines enable the comparison of vectors at different space-time points. For instance,  $\psi(x) - \psi(y)$  has no meaning since  $\psi(x)$  and  $\psi(y)$  live in different vector spaces attached to  $x$  and  $y$ , respectively.<sup>3</sup> To compare  $\psi(x)$  and  $\psi(y)$ , the Wilson line  $W[x, y]$  connecting  $x$  and  $y$  pulls back  $\psi(y)$  to  $x$  (or equivalently pushes forward  $\psi(x)$  to  $y$ )

$$\begin{aligned} W[x, y]\psi(y) - \psi(x), & \quad \text{to compare at } x; \\ \psi(y) - W[y, x]\psi(x), & \quad \text{to compare at } y. \end{aligned} \tag{2.84}$$

If comparison is allowed for vectors at different base points, the derivative associated to the Wilson line could be defined

$$[D_\mu \psi](x) = \lim_{a^\mu \rightarrow 0} \frac{W[x, x + a]\psi(x + a) - \psi(x)}{a^\mu}. \tag{2.85}$$

The derivative is covariant under some gauge transformation  $U(x)$ , i.e.,

$$[D_\mu U\psi](x) = U(x)[D_\mu \psi](x) \tag{2.86}$$

if and only if  $W[x, y]$  has the transformation property under  $U(x)$

$$W[x, y] \rightarrow U(x)W[x, y]U^{-1}(y). \tag{2.87}$$

This transformation rule is the reason why Wilson lines are so useful.

---

<sup>2</sup>For a quantum field  $\phi$ ,  $\phi(x)$  is a vector attached to  $x$  in the base space (space-time manifold). A quantum field configuration corresponds to a section of the bundle.

<sup>3</sup>Only when the connection is trivial is it meaningful to directly compare vectors at different points in the base space. Triviality means the following. Let  $(e_1, e_2, \dots)$  be a vector basis and  $\nabla_{e_i}$  be the covariant derivative along  $e_i$  associated with some connection. The connection is called trivial if and only if  $\nabla_{e_i} e_j = 0$  for any  $i$  and  $j$ .

Another equivalent way of defining the Wilson line is by the differential equation along a path  $x^\mu(s)$

$$\frac{dx^\mu}{ds} D_\mu W[x(s), x(s_0)] = 0, \quad (2.88)$$

with initial condition

$$W[x(s_0), x(s_0)] = 1. \quad (2.89)$$

To solve the differential equation Eq. (2.88), one can follow the regular procedure that is very familiar in time dependent perturbation theories. First define the following notation to be used shortly

$$V(s) \equiv ig \frac{dx^\mu}{ds} A_\mu(x(s)). \quad (2.90)$$

Then

$$\begin{aligned} & \frac{dx^\mu}{ds} D_\mu W[x(s), x(s_0)] = 0 \\ \Rightarrow & \left( \frac{d}{ds} - ig \frac{dx^\mu}{ds} A_\mu \right) W[x(s), x(s_0)] = 0 \\ \Rightarrow & W[x(s), x(s_0)] = 1 + \int_{s_0}^s ds_1 V(s_1) + \int_{s_0}^s ds_1 \int_{s_0}^{s_1} ds_2 V(s_1) V(s_2) + \cdots \end{aligned} \quad (2.91)$$

In general  $V(s_1)$  and  $V(s_2)$  might not commute, so

$$W[x(s), x(s_0)] = 1 + \int_{s_0}^s ds_1 V(s_1) + \frac{1}{2!} \mathbf{P} \int_{s_0}^s ds_1 \int_{s_0}^s ds_2 V(s_1) V(s_2) + \cdots, \quad (2.92)$$

where  $\mathbf{P}$  is the path order operator that orders the products of  $V(s_i)$ 's so that  $V(s_i)$  is always on the left side of  $V(s_j)$  if  $s_i > s_j$ . In a more compact fashion

$$W[x(s), x(s_0)] = \mathbf{P} \exp \left[ ig \int_{s_0}^s \frac{dx^\mu}{ds} A_\mu \right] \quad (2.93)$$

or

$$W[z, y] = \mathbf{P} \exp \left[ ig \int_y^z dx^\mu A_\mu \right] \quad (2.94)$$

where the integration is taken along some given path from  $y$  to  $z$ . For QED

$$W[z, y] = \exp \left[ -ie \int_y^z dx^\mu A_\mu \right], \quad (2.95)$$

and for QCD

$$W[z, y] = \mathbf{P} \exp \left[ ig_s \int_y^z dx^\mu A_\mu^a T^a \right]. \quad (2.96)$$

Following the defining equation (2.88), one immediately finds that under a gauge transformation  $U(x)$  the transformation rule Eq. (2.87) is true, by noticing that

$$\frac{dx^\mu}{ds} D'_\mu W'[x, y] = 0 \quad (2.97)$$

when plugging in

$$D'_\mu = U(x) D_\mu U^{-1}(x), \quad W'[x, y] = U(x) W[x, y] U^{-1}(y) \quad (2.98)$$

and using the uniqueness of the solution of the differential equation (2.88).

In SCET-I, there are two sorts of Wilson lines that are interesting. One of them is the ultra-soft Wilson line that describes soft radiation and the other one is useful for constructing collinear gauge invariant objects. Figure 6 shows the radiation of an ultra-soft gluon from a

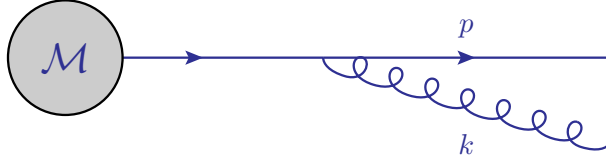


Figure 6: Radiation of an ultra-soft gluon from a collinear quark.

collinear quark. If the full QCD Feynman rule is used

$$\begin{aligned} \mathcal{M}_{\text{Fig. 6}} &= \bar{u}(p) i g_s \not{\epsilon}^*(k) \frac{i(\not{p} + \not{k} + m)}{(p+k)^2 - m^2} \mathcal{M} \\ &= -g_s \bar{u}(p) \frac{2p \cdot \epsilon^*(k) + \not{\epsilon}^*(k) \not{k}}{2p \cdot k} \mathcal{M}. \end{aligned} \quad (2.99)$$

In the limit where the gluon radiation is soft, i.e.,  $k \rightarrow 0$ ,

$$\mathcal{M}_{\text{Fig. 6}} = \bar{u}(p) \left[ -g_s \frac{p \cdot \epsilon^*(k)}{p \cdot k} \right] \mathcal{M}, \quad (2.100)$$

where  $\not{\epsilon}^* \equiv \epsilon^{*a} T^a$ . The factor (called the eikonal factor)

$$\left[ -g_s \frac{p \cdot \epsilon^*(k)}{p \cdot k} \right] \quad (2.101)$$

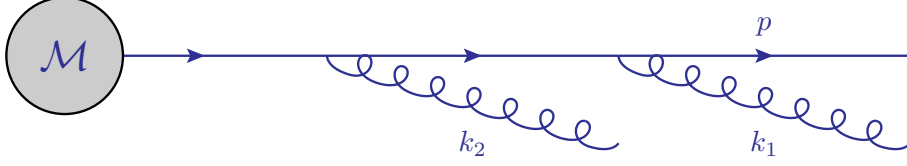


Figure 7: Radiation of two gluons from a collinear quark.

is quite universal. For two sequential ultra-soft gluon radiation as shown in Figure 7, the factor becomes

$$\left[ (-g_s)^2 \frac{p \cdot \epsilon^*(k_1)}{p \cdot k_1} \frac{p \cdot \epsilon^*(k_2)}{p \cdot (k_1 + k_2)} + (k_1 \leftrightarrow k_2) \right] \quad (2.102)$$

It turns out that Eq. (2.101) and (2.102) are parts of ultra-soft Wilson line in the momentum space. Let the direction of  $p$  in Figure 6 be  $n$ , then Eq. (2.101) becomes

$$\left[ -g_s \frac{n \cdot \epsilon^*(k)}{n \cdot k + i\epsilon} \right], \quad (2.103)$$

where it's important to put back the  $i\epsilon$  for the purpose below. The eikonal factor in Eq. (2.103) can be immediately obtained with the Lagrangian in Eq. (2.63), whose kinetic term is  $in \cdot \partial$  and interaction term with ultra-soft gluon is  $g_s \bar{\xi}_n n \cdot A_s \xi_n$ . Eq. (2.103) has a  $k$  in the denominator which corresponds to a derivative in the denominator. To deal with it, one can invoke a simple integral identity (here  $i\epsilon$  is essential)

$$\frac{1}{n \cdot k + i\epsilon} = -i \int_0^\infty ds e^{[is(n \cdot k + i\epsilon)]}. \quad (2.104)$$

Then we can rewrite Eq. (2.103) as

$$\begin{aligned} \left[ -g_s \frac{n \cdot \epsilon^*(k)}{n \cdot k + i\epsilon} \right] &= ig_s \int_0^\infty ds n \cdot \epsilon^*(k) e^{[is(n \cdot k + i\epsilon)]} \\ &= \langle k | ig_s \int_0^\infty ds n \cdot A(ns) e^{-\epsilon s} | 0 \rangle \\ &= \langle k | \mathbf{P} \exp \left[ ig_s \int_0^\infty ds n \cdot A(ns) e^{-\epsilon s} \right] | 0 \rangle \end{aligned} \quad (2.105)$$

where the gluon field  $A_\mu$ , like its momentum counter part  $\epsilon(k)$ , is short for  $A_\mu^a T^a$ . The remarkable thing is that this could be generalized to the radiation of an arbitrary number  $N$  of gluons

$$\langle k_1, k_2, \dots, k_N | \mathbf{P} \exp \left[ ig_s \int_0^\infty ds n \cdot A(ns) e^{-\epsilon s} \right] | 0 \rangle. \quad (2.106)$$

Denote

$$Y_n^\dagger(x) \equiv \mathbf{P} \exp \left[ ig_s \int_0^\infty ds n \cdot A(x + ns) e^{-\epsilon s} \right], \quad (2.107)$$

then  $Y_n^\dagger(0)$  describes ultra-soft radiation by an out-going particle along the  $n$ -direction. For incoming particle radiation, the ultra-soft Wilson line is

$$\tilde{Y}_n(x) \equiv \mathbf{P} \exp \left[ ig_s \int_{-\infty}^0 ds n \cdot A(x + ns) e^{+\epsilon s} \right], \quad (2.108)$$

and for out-going and in-coming anti-particles, the corresponding ultra-soft Wilson lines are  $Y_n$  and  $\tilde{Y}_n^\dagger$ , respectively.

Collinear Wilson lines are also useful for building gauge invariant operators. For instance, let

$$W_n(x) \equiv \mathbf{P} \exp \left[ ig_s \int_{-\infty}^0 ds \bar{n} \cdot A_n(x + \bar{n}s) e^{+\epsilon s} \right], \quad (2.109)$$

then

$$\chi_n(x) \equiv W_n^\dagger(x) \xi_n \quad (2.110)$$

and

$$\mathcal{A}^\mu \equiv W_n^\dagger(x) [i D_n^\mu W(x)] \quad (2.111)$$

are invariant under a collinear gauge transformation

$$U_n(x) \equiv \exp [i \alpha_n^a(x) T^a]. \quad (2.112)$$

These gauge invariant objects are very convenient as basic building blocks of the effective theories.



| I  | II   | III   |
|--|--|---|
| $n_\mu \rightarrow n_\mu + \Delta_\mu^\perp$ | $n_\mu \rightarrow n_\mu$                                  | $n_\mu \rightarrow e^\alpha n_\mu$                |
| $\bar{n}_\mu \rightarrow \bar{n}_\mu$        | $\bar{n}_\mu \rightarrow \bar{n}_\mu + \epsilon_\mu^\perp$ | $\bar{n}^\mu \rightarrow e^{-\alpha} \bar{n}^\mu$ |

Table 1: Three types of reparametrization invariance in SCET

### 2.1.6 Reparameterization Invariance

Reparameterization invariance (RPI) comes from the ambiguity of the way we label fields. So it is a redundancy our descriptions of the interaction similar to gauge invariance. Similar to gauge invariance, the requirement of RPI can impose constraints on possible interactions that can appear in the Lagrangian.

There are three types of RPI related to redefinition of the light-cone vectors  $n$  and  $\bar{n}$  that are shown in Table 1. The only restrictions of these redefinitions are that they must satisfy

$$n \cdot \bar{n} = 2, \quad n^2 = \bar{n}^2 = 0. \quad (2.113)$$

As an example, for type-I reparameterization,

$$(n_\mu + \delta_\mu)(n^\mu + \delta^\mu) = 0, \quad (n_\mu + \delta_\mu)\bar{n}^\mu = 2 \quad (2.114)$$

and Eq. (2.113) give rise to

$$\delta_\mu n^\mu = \delta_\mu \bar{n}^\mu = 0. \quad (2.115)$$

Ignoring  $\delta^2$  term and denoting  $\delta \rightarrow \Delta^\perp$  result in type-I transformation in Table 1. Note that in Table 1, type-I and type-II are infinitesimal transformations while type-III are finite.

These transformations should not upset the power counting of a typical momentum. For an  $n$ -collinear momentum,

$$n \cdot p \rightarrow (n + \Delta^\perp) \cdot p \sim \lambda^2 \quad (2.116)$$

implies  $\Delta^\perp \sim \lambda$ . Similary  $\epsilon^\perp \sim \lambda^0$  and  $\alpha \sim \lambda^0$  in Table 1.

We can also obtain the transformation properties of the fields. For instance, under a type-I transformation

$$\frac{\not{n}\not{\bar{n}}}{4}\xi_n = \xi_n \quad \Rightarrow \quad \delta_I \xi_n = \frac{\not{\Delta}^\perp \not{n}}{4}\xi_n. \quad (2.117)$$

RPI restricts the allowed forms of Lagrangians. One can show that the Lagrangian obtained in Eq. (2.63) is reparameterization invariant. It is also the unique Lagrangian at leading power that preserves Lorentz invariance, gauge invariance, reparameterization invariance.

### 2.1.7 Factorization Theorems

There are enormous applications of SCET to obtaining factorization theorems (applications in our work are presented in Chapter 3 and 4). In this section, we only discuss the factorization of the Sudakov factor in QCD which is in parallel with the discussion of that for  $\phi^3$  theory in Section 2.1.2.2.

To discuss the factorization of collinear and ultra-soft mode, the following decoupling transformation is essential [19]:

$$\begin{aligned} \xi_n(n) &= Y_n(x)\xi_n^{(0)}, \\ A_n^\mu(x) &= Y_n(x)A_n^{(0)\mu}Y_n^\dagger(x), \end{aligned} \quad (2.118)$$

where  $Y_n(x)$  is ultra-soft Wilson lines defined by

$$Y_n(x) \equiv \mathbf{P} \exp \left[ ig \int_{-\infty}^0 ds n \cdot A_{us}(x + ns) \right], \quad (2.119)$$

$\xi_n$  and  $A_n$  are collinear quark and gluon fields, respectively. Note that in the collinear Lagrangian Eq. (2.63), the only collinear and ultra-soft interaction comes from the term

$$\bar{\xi}_n i n \cdot D \frac{\not{n}}{2} \xi_n. \quad (2.120)$$

Using the defining identity of Wilson line as was discussed in detail in Section 2.1.5,

$$[n \cdot (\partial - ig_s A_{us}) Y_n(x)] = 0, \quad (2.121)$$

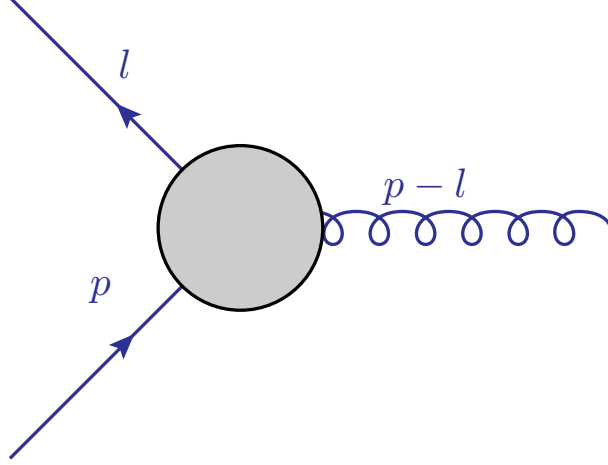


Figure 8: Sudakov form factor for QCD.

and applying the redefinition of fields in Eq. (2.118),

$$\begin{aligned}
\bar{\xi}_n i n \cdot D \frac{\not{n}}{2} \xi_n &= \bar{\xi}_n^{(0)} Y_n^\dagger i n \cdot [\partial - i g_s A_{us} - i g_s Y_n A_n^{(0)} Y_n^\dagger] \frac{\not{n}}{2} Y_n \xi_n^{(0)} \\
&= \bar{\xi}_n^{(0)} i n \cdot [\partial - i g_s A_n^{(0)}] \frac{\not{n}}{2} \xi_n^{(0)} \\
&= \bar{\xi}_n^{(0)} i n \cdot D_n^{(0)} \frac{\not{n}}{2} \xi_n^{(0)}
\end{aligned} \tag{2.122}$$

where

$$D_n^{(0)\mu} \equiv \partial^\mu - i g_s A_n^{(0)\mu}. \tag{2.123}$$

So with the decoupling transformation Eq. (2.118), the interaction between ultra-soft field  $A_{us}$  and quark collinear field  $\xi_n^{(0)}$  decouples. This, however, does not mean there is no interaction between the ultra-soft and collinear sector. We demonstrate this point below by studying the factorization of the Sudakov form factor in QCD.

To factorize the Sudakov form factor (corresponding to Figure 8), one needs to match the current operator between full QCD and SCET. For full QCD

$$J_{\text{QCD}}^\mu(x) = \bar{\psi}(x) \gamma^\mu \psi(x). \tag{2.124}$$

This operator should match to the SCET current operator

$$J_{\text{SCET}}^\mu(x) \sim \bar{\chi}_n(x) \gamma_\perp^\mu \chi_{\bar{n}}(x), \quad (2.125)$$

where  $\chi_n$  is the gauge invariant quark jet field defined in Eq. (2.110),  $\chi_{\bar{n}}$  is a quark jet field in the opposite direction of  $n$ , the reason for  $\sim$  instead of an equality is explained soon, and only  $\perp$  components of  $\gamma^\mu$  survive because

$$\frac{\not{n} \not{\bar{n}}}{4} \chi_n = \chi_n, \quad \bar{\chi}_n \frac{\not{\bar{n}} \not{n}}{4} = \bar{\chi}_n. \quad (2.126)$$

$J_{\text{SCET}}^\mu$  corresponds to, for instance, back-to-back two-jet productions. Similar to  $\phi^3$  theory, since  $\xi_n(x)$  is not power suppressed along  $\bar{n}$ , i.e.,

$$(\bar{n} \cdot \partial)^k \chi_n(x) \sim (\bar{n} \cdot p)^k \chi_n \sim \lambda^0 \chi_n, \quad (2.127)$$

one should include the operator  $\chi_n(x + \bar{s})$  for all  $s$  in the SCET current. Thus one should match  $J_{\text{QCD}}^\mu$  to

$$F^\mu(x) = \int ds \int ds' C(s, s') \bar{\chi}_n(x + \bar{n}s) \gamma_\perp^\mu \chi_{\bar{n}}(x + ns'), \quad (2.128)$$

where  $C(s, s')$  are Wilson coefficients that can be calculated by matching order by order from full QCD to SCET.

Now apply the decoupling transformation to  $F^\mu(x)$ :

$$\begin{aligned} F^\mu(x) &= \int ds \int ds' C(s, s') \bar{\chi}_n(x + \bar{n}s) \gamma_\perp^\mu \chi_{\bar{n}}(x + ns') \\ &= \int ds \int ds' C(s, s') \bar{\chi}_n^{(0)}(x + \bar{n}s) Y_n^\dagger(x + \bar{n}s) \gamma_\perp^\mu Y_{\bar{n}}(x + \bar{n}s) \chi_{\bar{n}}^{(0)}(x + ns'). \end{aligned} \quad (2.129)$$

Here is the key point. Since the interaction between  $\xi^{(0)}$  field and ultra-soft field  $A_{us}$  decouples, one obtains the following form of factorization<sup>4</sup>

$$|\langle p, l | F(0) | 0 \rangle|^2 = \int ds \int ds' |C(s, s', \mu)|^2 \mathcal{J}(s, s', \mu) \mathcal{S}(s, s', \mu), \quad (2.130)$$

---

<sup>4</sup>This formular is somewhat schetchy.  $F^\mu(x)$  could appear in different processes. For  $e^+e^-$  to 2 jets events, as an example, one would be interested in evaluating  $|\langle p, l | \epsilon(k^*)_\mu F^\mu(0) | 0 \rangle|^2$ , where  $\epsilon(k^*)$  is the off-shell photon (or  $Z$  bozon) polarization vector, and obtaining factorization formulae that look like the one in Eq. (2.130).

where  $\mathcal{J}$  comes from matrix elements involving quark jet fields  $\chi^{(0)}$  along both  $n$  and  $\bar{n}$  direction, and  $\mathcal{S}$  comes from matrix elements of ultra-soft Wilson lines along  $n$  and  $\bar{n}$  direction. So we see explicitly that the decoupling transformation does not get rid of interactions among different collinear sectors. Instead, the interactions manifest themselves through the connections to different collinear sectors by the ultra-soft fields  $A_{us}$ . Even though  $C$ ,  $\mathcal{J}$ , and  $\mathcal{S}$  appear in Eq. (2.130) with convolutions, the meaning of the factorization is that  $C$ ,  $\mathcal{J}$ , and  $\mathcal{S}$  have different characteristic scales. One can calculate different functions with different scales and use renormalization group techniques to resum large logarithms of ratios of different scales.

## 2.2 NON-RELATIVISTIC QCD

Non-relativistic QCD (NRQCD) is an effective field theory that was invented to describe interactions between heavy quarks at low energies, especially the dynamics of quarkonium bound states [26, 27, 28, 15]. Since relevant energy scales of NRQCD are very close to  $\Lambda_{QCD}$ , it pushes the applicability of the perturbative QCD to where it begins to breakdown.

### 2.2.1 NRQCD Power Counting Parameter $v$

Non-relativistic for NRQCD means that the velocity of heavy quarks under study is

$$\frac{v}{c} \ll 1. \quad (2.131)$$

To estimate the typical relative velocity of a quark inside a quarkonium (charmonium and bottomonium) state, one can invoke the virial theorem. The potential energy

$$V(r) \sim \frac{\alpha_s}{r}, \quad (2.132)$$

and the kinetical energy

$$T \sim \frac{1}{2} m_Q v^2. \quad (2.133)$$

|      | $c\bar{c}$ | $b\bar{b}$ | $t\bar{t}$ |
|------|------------|------------|------------|
| $M$  | 1.5 GeV    | 4.7 GeV    | 180 GeV    |
| $Mv$ | 0.9 GeV    | 1.5 GeV    | 16 GeV     |

Table 2: Quarkonium mass and momentum scales [1].

The balance between kinetic and potential energies (virial theorem) reads

$$2T = V(r) \quad \Rightarrow \quad m_Q v^2 \sim \frac{\alpha_s}{r}. \quad (2.134)$$

Using the uncertainty principle

$$\frac{1}{r} \sim p \sim m_Q v, \quad (2.135)$$

one obtains from Eq. (2.134) the estimation of  $v$

$$v \sim \alpha_s, \quad (2.136)$$

where  $\alpha_s$  should be evaluated at  $m_Q v$ . For charmonium, bottomonium, and toponium, the orders of  $m_Q v$  are shown in Table 2.

Some interesting scales involved in NRQCD include the heavy quark mass scale  $m_Q$ , the quark relative momentum inside a quarkonium  $m_Q v$ , the quark kinetic energy  $m_Q v^2$ , and  $\Lambda_{QCD}$ .

### 2.2.2 NRQCD Lagrangian

To obtain effective fields for heavy quarks, we work in the Dirac representation Eq. (2.48). This is convenient since below  $m_Q$  scale, heavy quark pair production is suppressed. The starting point is still full QCD

$$\mathcal{L}_Q = \bar{Q}(i\not{D} - m_Q)Q. \quad (2.137)$$

To integrate out the hard mode, introduce the following phase redefinition

$$Q(x) = e^{-im_Q t} \begin{pmatrix} \psi \\ \chi \end{pmatrix}, \quad (2.138)$$

which results in

$$\mathcal{L}(\psi, \chi) = \psi^\dagger iD^0 \psi + \chi^\dagger (iD^0 + 2m_Q) \chi - \psi^\dagger i\vec{\sigma} \cdot \vec{D} \chi - \chi^\dagger i\vec{\sigma} \cdot \vec{D} \psi. \quad (2.139)$$

The field redefinition in Eq. (2.138) effectively shifts mass from quark field  $\psi$  to the anti-quark partner. We can integrate out the massive mode using equation of motion for  $\chi$

$$\begin{aligned} (iD^0 + 2m_Q) \chi &= i\vec{\sigma} \cdot \vec{D} \psi \\ \Rightarrow \quad \chi &= \frac{1}{iD^0 + 2m_Q} i\vec{\sigma} \cdot \vec{D} \psi, \end{aligned} \quad (2.140)$$

and find the Lagrangian for quark field  $\psi$

$$\mathcal{L}(\psi) = \psi^\dagger iD^0 \psi - \psi^\dagger i\vec{\sigma} \cdot \vec{D} \frac{1}{iD^0 + 2m_Q} i\vec{\sigma} \cdot \vec{D} \psi \quad (2.141)$$

Expanding in powers of  $m_Q^{-1}$

$$\mathcal{L}(\psi) = \psi^\dagger (iD^0 + \frac{\vec{D}^2}{2m_Q}) \psi + \frac{g_s}{2m_Q} \psi^\dagger \vec{\sigma} \cdot \vec{B} \psi + \mathcal{O}(\frac{1}{m_Q^2}), \quad (2.142)$$

where

$$B_i = \frac{1}{2} \epsilon_{ijk} F^{jk}, \quad \text{with} \quad F^{\mu\nu} \equiv \frac{i}{g} [D^\mu, D^\nu]. \quad (2.143)$$

One can do a similar field redefinition to integrate out particle mode

$$Q(x) = e^{im_Q t} \begin{pmatrix} \psi \\ \chi \end{pmatrix}, \quad (2.144)$$

from which the effective Lagrangian for  $\chi$  is

$$\mathcal{L}(\chi) = \chi^\dagger (iD^0 - \frac{\vec{D}^2}{2m_Q}) \chi - \frac{g_s}{2m_Q} \chi^\dagger \vec{\sigma} \cdot \vec{B} \chi + \mathcal{O}(\frac{1}{m_Q^2}). \quad (2.145)$$

Note that the NRQCD Lagrangian allows both  $\mathcal{L}(\psi)$  and  $\mathcal{L}(\chi)$  for quark and anti-quark fields, respectively, but does not allow heavy quark pair production.

The next step is to figure out the power (velocity  $v$ ) counting for heavy quark and gluon fields. Let the normalization of quarkonium bound state be

$$\langle H | H \rangle = 1. \quad (2.146)$$

The size of the bound states is of the order

$$\int d^3x \sim r^3 \sim \frac{1}{(m_Q v)^3}, \quad (2.147)$$

and the quark number operator is

$$\hat{N} = \int d^3x \psi^\dagger \psi. \quad (2.148)$$

So from

$$\langle H | \hat{N} | H \rangle \sim 1 \quad (2.149)$$

one can get the power counting of  $\psi$  as

$$\psi \sim (m_Q v)^{3/2}. \quad (2.150)$$

The power counting of  $\vec{D}$  is associated with momentum  $\vec{p}$ , so  $\vec{D} \sim m_Q v$ . Another way of seeing this is to look at the kinetic energy operator

$$\text{KineticEnergy} = \int d^3x \psi^\dagger(x) \frac{\vec{D}^2}{2m_Q} \psi(x) \sim m_Q v^2. \quad (2.151)$$

| Operator                         | Power counting  |
|----------------------------------|-----------------|
| $\psi$                           | $(m_Q v)^{3/2}$ |
| $\chi$                           | $(m_Q v)^{3/2}$ |
| $D_0$                            | $m_Q v^2$       |
| $\vec{D}$                        | $m_Q v$         |
| $g_s \vec{E}$                    | $m_Q^2 v^3$     |
| $g_s \vec{B}$                    | $m_Q^2 v^4$     |
| $g_s A_0$ (in Coulomb gauge)     | $m_Q v^2$       |
| $g_s \vec{A}$ (in Coulomb gauge) | $m_Q v^3$       |

Table 3: Power counting rules for basic NRQCD operator building blocks [1].



This, combining with Eq. (2.147) and (2.150), also gives  $\vec{D} \sim m_Q v$ . Using the equation of motion

$$(iD^0 + \frac{\vec{D}^2}{2m_Q})\psi(x) = 0 \quad (2.152)$$

one finds  $D^0 \sim m_Q v^2$ .

By this sort of analysis, one can obtain a list of power counting rules for all basic operators which are shown in Table 3. With this list one can build more complex operators in the NRQCD Lagrangian and other composite operators. Note that the power counting of  $A_\mu$  in Table 3 is dependent on the gauge choice because gauge dependent equation of motion for  $A_\mu$  is used for its estimation.

### 2.2.3 Quarkonium Production and NRQCD Factorization Formalism

The NRQCD factorization formalism works for sufficiently inclusive processes and factorizes relevant observables into perturbative short distance coefficients that are perturbatively calculable order by order and non-perturbative long distance matrix elements (LDMEs) [27]. LDMEs are supposed to be universal and should be extracted from data. In Chapter 4, we study quarkonium production in jets which on the one hand uses the factorization formalism from SCET and on the other hand uses the NRQCD factorization formalism. In particular, we use the factorized fragmentation functions obtained from NRQCD which describe physics below heavy quark mass scale  $m_Q$ .

For quarkonium production in the NRQCD factorization formalism, the short distance coefficients contain all physics of scales  $\gtrsim m_Q$ , i.e., they account for the production of an quark anti-quark pair. The physics for scales  $\lesssim (m_Q v)$  are described by the NRQCD Lagrangian discussed in Section 2.2.2. So, pictorially, a  $Q\bar{Q}$  pair is produced in a space-time point and this pair evolves to a heavy quark bound state. This evolution is described by the production operators [27]:

$$\begin{aligned} \mathcal{O}_n^H &= \chi^\dagger \mathcal{K}'_n \psi \left( \sum_X \sum_{m_J} |H+X\rangle \langle H+X| \right) \psi^\dagger \mathcal{K}_n \chi \\ &= \chi^\dagger \mathcal{K}'_n \psi \left( a_H^\dagger a_H \right) \psi^\dagger \mathcal{K}_n \chi, \end{aligned} \quad (2.153)$$

where  $H$  is the final quarkonium bound state,  $X$  includes all final particles except for  $H$ , the dummy index  $m_J$  sums over the  $2J+1$  spin states of  $H$ ,  $\psi^\dagger \mathcal{K}_n \chi$  create a  $Q\bar{Q}$  pair with different transformation properties depending on the factor  $\mathcal{K}_n$  chosen.  $\mathcal{K}_n$  could be combinations of spin matrices  $\sigma^i$ , color matrices  $T^a$ , and covariant derivatives  $\vec{D}$ . For instance, if  $\mathcal{K}_n = 1$ , a  $Q\bar{Q}$  pair is created by  $\psi^\dagger \mathcal{K}_n \chi$  that transforms as a spin and color singlet which is usually denoted as  $^1S_0^{[1]}$ . Or if  $\mathcal{K}_n = \sigma^i T^a$ , a  $Q\bar{Q}$  pair is created by  $\psi^\dagger \mathcal{K}_n \chi$  that transforms as a spin triplet and color octet which is usually denoted as  $^3S_1^{[8]}$ .

The NRQCD factorization approach separates cross sections into short distance pieces and LDMEs:

$$\sigma(H) = \sum_n d_n \langle 0 | \mathcal{O}_n^H | 0 \rangle, \quad (2.154)$$

where  $d_n$  are short distance coefficients (with scale  $\geq m_H$ ) that are calculable perturbatively depending on the process in consideration, and  $\langle 0 | \mathcal{O}_n^H | 0 \rangle$  are nonperturbative LDMEs that describe evolution of quark anti-quark pairs into quarkonium bound states  $H$ . The relative importance of different LDMEs can be estimated by power counting  $\mathcal{O}_n^H$  in terms of relative velocity  $v$  of quarks in the quarkonium center of mass frame based on velocity counting rules shown in Table 3. Similarly, quarkonium fragmentation functions, which are the probabilities that a parton decays into a quarkonium and anything else, factorize as follows:

$$D_{q/g}^H = \sum_n d_{q/g,n} \langle 0 | \mathcal{O}_n^H | 0 \rangle, \quad (2.155)$$

with  $d_{q/g,n}$  calculable short distance coefficients and the LDMEs

$$\langle 0 | \mathcal{O}_n^H | 0 \rangle \quad (2.156)$$

are the same LDMEs as those in Eq. (2.154).

The factorized formulae of the form shown in Eq. (2.155) will be used in Chapter 4 for studying quarkonium production in jets. Such analysis in turn serves as a test of the NRQCD factorization formalism and the universalities of LDMEs. As an example of Eq. (2.155) [29, 30], the fragmentation function of  $J/\psi$  from a gluon with state  $^1S_0^{[8]}$  has the following expression:

$$D_g^{^1S_0^{[8]}}(z, 2m_c) = \frac{5\alpha_s^2(2m_c)}{96m_c^3} \langle \mathcal{O}^\psi(^1S_0^{[8]}) \rangle (3z - 2z^2 + 2(1-z)\ln(1-z)), \quad (2.157)$$

where the LDME  $\langle \mathcal{O}^\psi(^1S_0^{[8]}) \rangle$  is non-perturbative and the rest is the perturbatively calculable coefficient.

## 3.0 FRAGMENTATION TO A JET

This chapter is devoted to the study of jets. Sections 3.1 and 3.2 are based on our work in Refs. [31] and [32], respectively. Each section is self-contained and can be read independently.<sup>1</sup>

In Section 3.1, a function called fragmentation function to a jet (FFJ) is introduced to describe the inclusive production of a jet from a parton. We use FFJ to study the properties of jets with small radii. In Section 3.2, FFJ formalism is used further to study the production of jets in the large  $z$  limit where  $\ln(1 - z)$  is resummed.

## 3.1 FRAGMENTATION OF A JET WITH SMALL RADIUS

### 3.1.1 Introduction

Jets in high energy collisions have been an important theoretical and experimental probe of physics for decades. Currently, they are not only important for understanding Quantum Chromodynamics (QCD), but are crucial in our searches for beyond the Standard Model physics at the Large Hadron Collider at CERN and will continue to be important for any future collider that may be built. Understanding the property of jets and being able to calculate reliable cross sections to compare to data are thus extremely important to current and future studies in particle physics. Current understanding of jets relies on Monte Carlo simulations and effective field theories, both of them have been developed very fast during

---

<sup>1</sup>To avoid confusion, in this chapter (the same in the next chapter), the reference “section” will specifically have two index numbers such as Section 3.1 and Section 3.2 while “subsection” has more than two index numbers such as Subsection 3.1.4 and Subsection 3.2.3.1.

the past two decades and the work of either of them is far from being finished.

Conceptually, a jet is a collinear set of energetic particles in the detector<sup>2</sup>. In order to make this concept concrete, there needs to be some jet algorithm to define how particles are sorted to be within or outside of the jet. Most jet algorithms use a parameter to differentiate the two sets of particles, often denoted as the jet radius  $R$ . When doing theoretical calculations involving a jet definition, logarithms of this new object occur,  $\ln R$ , and thus to make sure that we have perturbative convergence of QCD, choosing  $R \sim 1$  would be natural. However, it is sometime useful to investigate narrow jets by choosing a smaller  $R$ , since it can help resolve individual jets, remove pileup, and probe jet substructure. This leads to the problem of the breakdown of perturbation theory, and requires resummation of  $\ln R$ . This has been investigated in QCD in Refs. [33, 34, 35, 36]. Since jets are made up of collinear particles, soft-collinear effective theory (SCET) [16, 17, 19, 20] is a natural tool to study jets. Indeed, there have been many studies of  $\ln R$  resummation within SCET [37, 38, 39, 40].

In this section, we introduce the fragmentation function to a jet (FFJ) in SCET. The FFJ,  $D_{J_k/l}(z, \mu)$ , describes the fragmentation of parton  $l$  into a jet with momentum fraction  $z$  containing parton  $k$ . We calculate the different possible combinations of quark and gluon initial and final partons. By summing over final state partons, we obtain the inclusive FFJ,  $D_{J/l}(z, \mu)$ , describing the inclusive fragmentation of parton  $l$  into a jet. The renormalization of this object will be shown to lead to the standard Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution with the natural scale dependent on  $R$ , and thus we can use this object and the renormalization group to resum logarithms of  $R$ .

We also present a new factorization theorem for the fragmentation of a hadron within a jet, where the FFJ appears, allowing for the resummation of  $\ln R$  for this process. We further generalize this factorization for the situation of a subjet with radius  $r$  within a fat jet of radius  $R$ . This allows the resummation of the ratio of these radii,  $\ln R/r$ .

The organization of this section is as follows. In Subsection 3.1.2, we give the definition of the FFJ in SCET, and calculate next-to-leading (NLO) corrections. From this we can derive the renormalization group behavior and see that it is the standard DGLAP evolution.

---

<sup>2</sup>depending on the physical processes and jet algorithms, a jet can have an arbitrary number of particles (including the special case of just having one particle).

In Subsection 3.1.3 we present the factorization theorem for the fragmentation inside a jet. Combining this with the renormalization behavior from the previous subsection allows for resummation of  $\ln R$  for this process. In Subsection 3.1.4 we consider the subjet fragmentation from a fat jet. We conclude in Subsection 3.1.5. Finally, in Appendix A.1 we describe hadron fragmentation inside of a jet, which is very similar to the subjet fragmentation of Subsection 3.1.4.

*Note:* While completing this work, Ref. [41] appeared on the arXiv with significant overlap with Subsection 3.1.2.

### 3.1.2 Inclusive Jet Fragmentation Function

The definition of the fragmentation function to a jet (FFJ) is similar to the fragmentation function to a hadron (HFF). In SCET, if a collinear quark,  $q$ , fragments to a jet with a momentum fraction  $z$ , the probability is given as

$$\begin{aligned}
& D_{J_k/q}(z, \mu) \\
&= \sum_{X_{\not{J}}, X_{J-1}} \frac{1}{2N_c z} \int d^{D-2} \mathbf{p}_J^\perp \text{Tr} \langle 0 | \delta \left( \frac{p_J^+}{z} - \mathcal{P}_+ \right) \delta^{(D-2)}(\mathcal{P}_\perp) \frac{\bar{\not{p}}}{2} \Psi_n | J_k(p_J^+, \mathbf{p}_J^\perp, R) X_{\not{J}} \rangle \\
&\quad \times \langle J_k(p_J^+, \mathbf{p}_J^\perp, R) X_{\not{J}} | \bar{\Psi}_n | 0 \rangle,
\end{aligned} \tag{3.1}$$

where  $\Psi_n = W_n^\dagger \xi_n$ ,  $W_n$  is a collinear Wilson line in SCET [17, 19], and  $R$  is the jet radius to be determined by specific jet algorithm.  $X_{J-1}$  are the final states included in the observed jet except the primary jet parton  $k$  and  $X_{\not{J}}$  are final states not included in the jet. Throughout, we will work in  $D = 4 - 2\varepsilon$  dimensions, and use the convention,  $p_+ \equiv \bar{n} \cdot p = p_0 + \hat{\mathbf{n}}_J \cdot \mathbf{p}$ ,  $p_- \equiv n \cdot p = p_0 - \hat{\mathbf{n}}_J \cdot \mathbf{p}$ , where  $\hat{\mathbf{n}}_J$  is a unit vector in the jet direction. The lightcone vectors  $n$  and  $\bar{n}$  satisfy  $n^2 = \bar{n}^2 = 0$  and  $n \cdot \bar{n} = 2$ . Therefore  $p_+ \sim 2E$  for a collinear particle in  $\hat{\mathbf{n}}_J$  direction. The expression of FFJ in Eq. (3.1) is displayed in the parton frame, where the transverse momentum of the mother parton,  $\mathbf{p}_\perp$ , is zero.

If we consider FFJ in the jet frame, where the transverse momentum of the observed jet,  $\mathbf{p}_J^\perp = 0$ , we can do the integral on  $\mathbf{p}_J^\perp$  using the relation  $\mathbf{p}_\perp = -\mathbf{p}_J^\perp/z$ . As a result we can

express FFJ as

$$D_{J_k/q}(z, \mu) = \sum_{X_{\notin J}, X_{J-1}} \frac{z^{D-3}}{2N_c} \text{Tr} \langle 0 | \delta \left( \frac{p_J^+}{z} - \mathcal{P}_+ \right) \frac{\vec{\not{p}}}{2} \Psi_n | J_k(p_J^+, R) X_{\notin J} \rangle \langle J_k(p_J^+, R) X_{\notin J} | \bar{\Psi}_n | 0 \rangle. \quad (3.2)$$

The normalization is chosen so that at lowest order (LO) in  $\alpha_s$ , the FFJ is given by

$$D_{J_q/q}^{(0)}(z) = \frac{z^{D-3}}{2N_c} \text{Tr} \frac{\vec{\not{p}}}{2} p_J^+ \frac{\not{p}}{2} \delta \left( \frac{p_J^+}{z} - p_J^+ \right) \cdot N_c = \delta(1 - z). \quad (3.3)$$

Like usual fragmentation functions to hadrons (HFFs), the FFJ satisfies the following momentum conservation,

$$\sum_{k=q, \bar{q}, g} \int_0^1 dz z D_{J_k/q}(z, \mu) = 1. \quad (3.4)$$

When a gluon initiates a jet fragmentation, the gluon FFJ in the parton frame is defined as

$$D_{J_k/g}(z, \mu) = \sum_{X_{\notin J}, X_{J-1}} \frac{1}{p_J^+(D-2)(N_c^2-1)} \int d^{D-2} \mathbf{p}_J^\perp \times \text{Tr} \langle 0 | \delta \left( \frac{P_+}{z} - \mathcal{P}_+ \right) \delta^{(D-2)}(\mathbf{P}_\perp) \mathcal{B}_n^{\perp\mu, a} | J_k(p_J^+, \mathbf{p}_J^\perp, R) X_{\notin J} \rangle \langle J_k(p_J^+, \mathbf{p}_J^\perp, R) X_{\notin J} | \mathcal{B}_{n\mu}^{\perp a} | 0 \rangle. \quad (3.5)$$

Here  $\mathcal{B}_n^{\perp a}$  is a covariant collinear gluon field strength, defined by  $\mathcal{B}_n^{\perp\mu, a} = i\bar{n}^\rho g_\perp^{\mu\nu} G_{n, \rho\nu}^a \mathcal{W}_n^{ba} = i\bar{n}^\rho g_\perp^{\mu\nu} \mathcal{W}_n^{\dagger, ba} G_{n, \rho\nu}^b$ , where  $\mathcal{W}_n$  is the collinear Wilson line in the adjoint representation. It satisfies

$$\mathcal{B}_n^{\perp\mu} = \mathcal{B}_n^{\perp\mu, a} T^a = \frac{1}{g} W_n^\dagger \left[ \bar{n} \cdot iD_n, iD_n^{\perp\mu} \right] W_n = \frac{1}{g} \left[ \mathcal{P}_+ W_n^\dagger iD_n^{\perp\mu} W_n \right]. \quad (3.6)$$

For defining the jet, we will employ an inclusive  $k_T$ -type algorithm. This is a recombination algorithm, which has the same constraint for  $k_T$  [42, 43], C/A [44], and anti- $k_T$  [45] up to NLO in  $\alpha_s$ . If two particles merge into a jet, the constraint is given by

$$\theta < R \quad (e^+e^- \text{ collider}), \quad (3.7)$$

$$\theta < \frac{R}{\cosh y} \quad (\text{hadron collider}), \quad (3.8)$$

where  $\theta$  is the angle between two particles, and  $y$  is the rapidity that describes the boost of the jet along the beam axis. For a hadron collider, we assumed  $\Delta y$  and  $\Delta\phi$  are small, so Eq. (3.8) is applicable to the jet with small  $R$ . When we compute NLO corrections to the jet

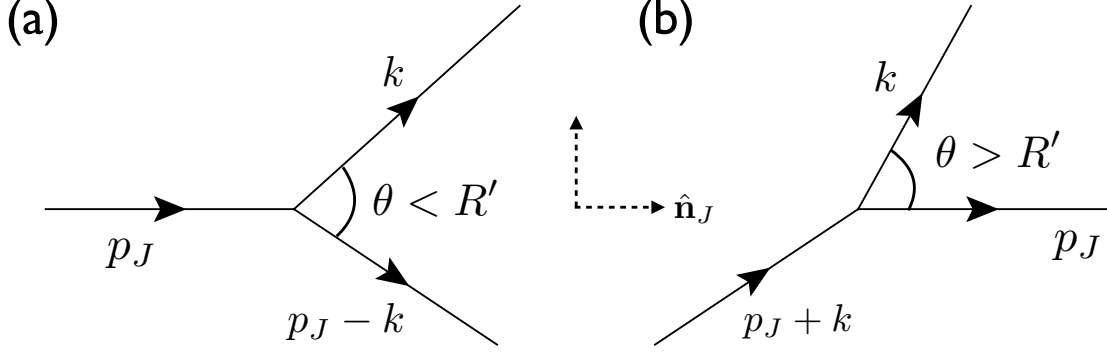


Figure 9: Jet fragmentation at NLO in  $\alpha_s$ . Diagram (a) shows the jet merging, so the contribution to FFJ should be proportional to  $\delta(1-z)$ . Diagram (b) shows the jet splitting, which has a contribution with a fraction  $z < 1$ .

algorithm, we will use  $\theta < R'$  for the sake of simplicity, where  $R' = R$  for  $e^+e^-$  colliders and  $R' = R/\cosh y$  for hadron colliders. As we will see later, typical scales for jet functions are  $p_+ \tan(R'/2)$ . In the small  $R$  limit,  $p_+ \tan(R'/2) \sim ER'$  are approximated as  $ER$  for  $e^+e^-$  annihilation and  $p_T R$  for hadron collision, where  $p_T$  is the transverse momentum of the jet to the hadron beam direction.

Fig. 9 shows the two possible cases for jet fragmentation. If  $\theta < R'$ , shown in Fig. 9-(a), the two particles in the final states are combined into a jet and the jet fraction is proportional to  $\delta(1-z)$ . In this case the phase space constraint in the jet frame ( $\mathbf{p}_J^\perp = 0$ ) is given by [46]

$$\tan^2 \frac{R'}{2} > \frac{p_J^{+2} k_-}{(p_J^+ - k_+)^2 k_+} . \quad (3.9)$$

If  $\theta > R'$ , only one particle is chosen to be in the jet, shown in Fig. 9-(b), hence the jet splitting arises with the fraction  $z$ . The phase space constraint in the jet frame becomes

$$\tan^2 \frac{R'}{2} < \frac{k_-}{k_+} . \quad (3.10)$$

There appears to be a gap in the phase space between Eqs. (3.9) and (3.10). However when we express the momentum of the mother parton as  $p$ , we have  $\mathbf{p}_\perp = \mathbf{p}_J^\perp = 0$  for Eq. (3.9) but  $\mathbf{p}_\perp = \mathbf{k}_\perp$  for Eq. (3.10) with  $p_J = p - k$ . Therefore when we express  $k_-$  in



terms of  $p_+$  and  $p^2$ ,  $k_-$  is different in Eq. (3.9) and (3.10);  $k_- = (1-x)p^2/p_+$  for Eq. (3.9) and  $k_- = p^2/((1-x)p_+)$  for Eq. (3.10), where  $x = k_+/p_+$ . So the right sides of the inequalities Eqs. (3.9) and (3.10) end up both equaling  $p^2/(x(1-x)p_+^2)$  and there is no gap in phase space.

**3.1.2.1 NLO Calculation of Quark FFJ** Following the description in Fig. 9, it is convenient to separate the full NLO contribution into ‘jet merging’ ( $\theta < R'$ ) and ‘jet splitting’ ( $\theta > R'$ ) contributions. In the jet merging contribution, the momentum of the mother parton is equal to the jet momentum,  $p_J$ . For quark initiated jets, it can be described by

$$D_{J/q}^{\text{in}}(z; E_J R') = \delta(1-z) \int_0^{\Lambda^2} dM^2 \frac{1}{2N_c p_J^+} \quad (3.11)$$

$$\times \sum_{X_{J-1}} \text{Tr} \langle 0 | \delta(M^2 - \mathcal{P}^2) \frac{\not{n}}{2} \Psi_n | J_q(p_J^+, R) \rangle \langle J_q(p_J^+, R) | \bar{\Psi}_n | 0 \rangle,$$

where  $M^2$  is the invariant mass of the final states. The gluon case is similarly defined with  $\mathcal{B}_n^{\perp\mu,a}$ .  $\Lambda^2$  is the maximal jet mass when  $\theta = R'$ . As there are two particles in the final state,  $\Lambda^2$  is usually also dependent of each particle’s energy. This jet merging contribution includes all the virtual corrections. Therefore combining the real and virtual contributions we can cancel all the infrared (IR) divergences and the result has only ultraviolet (UV) divergences.

Note that other than the  $\delta(1-z)$ , Eq. (3.11) is closely related to the standard quark jet function in SCET, defined as

$$\sum_{X_n} \langle 0 | \Psi_n^\alpha | X_n \rangle \langle X_n | \bar{\Psi}_n^\beta | 0 \rangle = \int \frac{d^4 p_{X_n}}{(2\pi)^3} p_{X_n}^+ \frac{\not{n}}{2} J_q(p_{X_n}^2) \delta^{\alpha\beta}. \quad (3.12)$$

Here  $J_q$  is normalized as  $J_q^{(0)}(p^2) = \delta(p^2)$  at LO in  $\alpha_s$ . Using this, we can rewrite Eq. (3.11) to be

$$D_{J/q}^{\text{in}}(z; E_J R') = \delta(1-z) \int_0^{\Lambda^2} dM^2 J_q(M^2; \theta < R') = \delta(1-z) \mathcal{J}_q(E_J R'; \theta < R'), \quad (3.13)$$

where  $J_q(M^2; \theta < R')$  is the unintegrated jet function for the final states inside the jet and  $\mathcal{J}_q$  is the integrated jet function (also called the unmeasured jet function in Ref [46]). Both have been computed to NLO in Ref. [46, 47, 48] with  $k_T$ -type and cone-type algorithms

applied. When we apply the  $k_T$ -type algorithm in Eq. (3.7), the jet merging contribution to NLO is given by

$$D_{J/q}^{\text{in}}(z; E_J R') = \delta(1-z) \left\{ 1 + \frac{\alpha_s C_F}{2\pi} \left[ \frac{1}{\varepsilon_{\text{UV}}^2} + \frac{1}{\varepsilon_{\text{UV}}} \left( \frac{3}{2} + \ln \frac{\mu^2}{p_J^{+2} t^2} \right) + \frac{3}{2} \ln \frac{\mu^2}{p_J^{+2} t^2} + \frac{1}{2} \ln^2 \frac{\mu^2}{p_J^{+2} t^2} + \frac{13}{2} - \frac{3\pi^2}{4} \right] \right\}, \quad (3.14)$$

where  $t \equiv \tan(R'/2) \sim R'/2$ .

Note that the renormalization behavior of the unintegrated jet function  $J_q(M^2; \theta < R')$  in Eq. (3.13) is different from the standard jet function without the restriction in Eq. (3.12). For example, all the UV divergences of the unintegrated jet function are only proportional to  $\delta(M^2)$  while this is not true for the standard jet function. The main reason for this difference comes from different treatments of the zero-bin subtraction [49]. For the unintegrated jet function in the small  $R$  limit, the relevant zero-bin subtracted mode should specifically be a collinear-soft mode [37, 38, 50] with scaling  $(p_{cs}^+, p_{cs}^\perp, p_{cs}^-) \sim Q\eta(1, R, R^2)$ , where  $\eta$  is a small parameter. This mode can resolve the jet boundary. Since the contribution of this collinear-soft mode to the jet mass squared is much smaller than  $M^2 \sim E_J^2 R'^2$ , UV divergences coming from this mode's zero-bin subtraction only contribute to the  $\delta(M^2)$  part. The details of the computation with this collinear-soft mode have been shown in Ref. [48]. However, in case of the standard jet function, the zero-bin subtracted mode is an ordinary soft mode and its contribution to the jet mass is non-negligible. For this type of zero-bin subtraction we obtain UV divergences proportional to  $1/M^2$  as well as  $\delta(M^2)$ .

In addition, there have been some complications about the integrability relation between the unintegrated and the integrated jet functions in Eq. (3.13). When  $M^2 \sim E_J^2 R'^2$  in the small  $R$  limit as considered in this section, we can describe the unintegrated jet function using only the collinear mode scaling as  $(p_+, p_\perp, p_-) \sim Q(1, R, R^2)$ , resulting in the integrability relation in Eq. (3.13). However in case of  $M^2 \ll E_J^2 R'^2$ , the integrated jet function is obtained from the convolution of the standard-like jet function and the soft function [46], where the standard-like jet function has the same UV behavior as the standard jet function. A concrete discussion about these differences can be found in Ref. [38].

For the jet splitting contribution, at least one particle in the final state should not be included in the jet. It therefore can be written as

$$D_{J_{k/q}}^{\text{out}}(z; p_+ R'/2) = \int_{\Lambda^2}^{\infty} dM^2 \frac{z^{D-3}}{2N_c} \times \sum_{X \notin J} \text{Tr} \langle 0 | \delta\left(\frac{p_J^+}{z} - \mathcal{P}_+\right) \delta(M^2 - \mathcal{P}^2) \frac{\vec{\not{p}}}{2} \Psi_n | J_k X_{\notin J} \rangle \langle J_k X_{\notin J} | \bar{\Psi}_n | 0 \rangle, \quad (3.15)$$

where  $p_+ = p_J^+/z \sim 2E = 2E_J/z$  is two times of mother parton's energy. At NLO we can have at most two particles in the final state, so we can further separate this contribution as quark or gluon jet contributions. For the quark jet contribution, the gluon should be outside the jet, and vice versa for the gluon jet.

First let us consider the quark jet contribution, where the momentum of the final state quark is given by  $p_J$ . In this case the gluon outside the jet becomes soft as  $z$  goes to 1, leading to an IR singularity in the naive collinear computation unless we subtract the zero-bin contribution [49]. In order to isolate the singularity as  $z \rightarrow 1$ , we can write the quark jet contribution as follows:

$$D_{J_{q/q}}^{\text{out}}(z; ER') = \delta(1-z) \left( \int_0^1 dz' D_{\text{out}}^q(z'; E_J R') \right) + \left[ D_{\text{out}}^q(z; ER') \right]_+. \quad (3.16)$$

Here the second term follows the standard plus distribution and is free of IR divergences as  $z \rightarrow 1$ .

In Fig. 10 we show the quark jet splitting contributions diagrammatically where the gluon in the final state cannot be merged into a quark jet with momentum  $p_J$ . The contribution of Fig. 10-(a) is given by

$$\begin{aligned} D_{J_{q/q}}^{\text{out},(a)} &= 4\pi g^2 C_F \mu_{\overline{\text{MS}}}^{2\varepsilon} \int_{\Lambda^2}^{\infty} dM^2 \frac{p_J^+}{M^2} \frac{z^{D-3}}{1-z} \int \frac{d^D k}{(2\pi)^D} \delta(k^2) \delta\left(\frac{1-z}{z} p_J^+ - k_+\right) \delta(M^2 - p_J^+ k_-), \\ &= \frac{\alpha_s C_F}{2\pi} \frac{(\mu^2 e^\gamma)^\varepsilon}{\Gamma(1-\varepsilon)} \int_{\Lambda^2}^{\infty} \frac{dM^2}{(M^2)^{1+\varepsilon}} z^{1-\varepsilon} (1-z)^{-1-\varepsilon}, \end{aligned} \quad (3.17)$$

where  $\mu_{\overline{\text{MS}}}^2 = \mu^2 e^\gamma / (4\pi)$ , and  $\Lambda$  is the maximal jet mass for  $\theta = R'$ ,

$$\Lambda^2 = p_J^{+2} t^2 \frac{1-z}{z} = p_+^2 t^2 z(1-z). \quad (3.18)$$

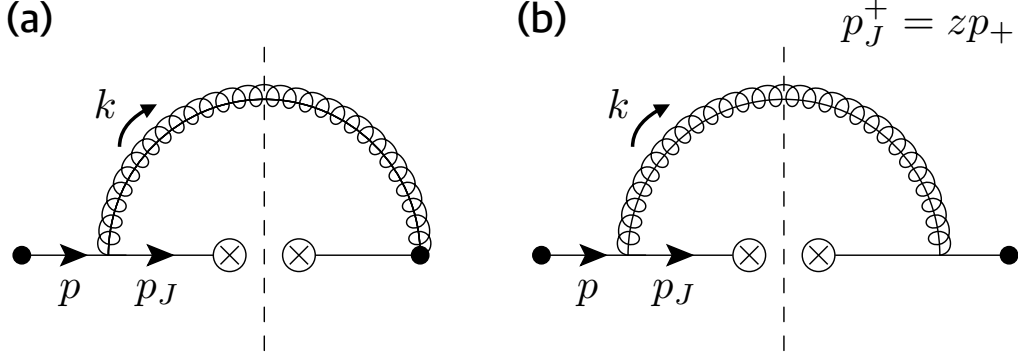


Figure 10: Feynman diagrams for quark jet splitting contribution at NLO in  $\alpha_s$ . Here the dashed lines represent the unitary cuts. The gluon in the final state is outside the jet. Diagram (a) has its Hermitian conjugate contribution.

As  $z$  become close to 1, Eq. (3.17) has IR divergence arising from soft gluon radiation. It is cancelled by the subtraction of the zero-bin contributions. The diagram Fig. 10-(b) gives

$$\begin{aligned}
D_{J_q/q}^{\text{out},(b)} &= 4\pi g^2 C_F \mu_{\overline{\text{MS}}}^{2\varepsilon} (1 - \varepsilon) \int_{\Lambda^2}^\infty dM^2 \frac{z^{D-3} k_+}{M^2} \int \frac{d^D k}{(2\pi)^D} \delta(k^2) \delta\left(\frac{1-z}{z} p_J^+ - k_+\right) \delta(M^2 - p_J^+ k_-) \\
&= \frac{\alpha_s C_F}{2\pi} \frac{(\mu^2 e^\gamma)^\varepsilon}{\Gamma(1-\varepsilon)} (1 - \varepsilon) \int_{\Lambda^2}^\infty \frac{dM^2}{(M^2)^{1+\varepsilon}} z^{-\varepsilon} (1-z)^{1-\varepsilon}.
\end{aligned} \tag{3.19}$$

Including the hermitian conjugate of diagram Fig. 10-(a), the final result for the jet splitting is  $D_{\text{out}}^q = 2D_{\text{out}}^{q,(a)} + D_{\text{out}}^{q,(b)}$ .

To calculate the part of Eq. (3.16) proportional to  $\delta(1-z)$ , we integrate over  $z$ ,

$$\begin{aligned}
\int_0^1 dz D_{\text{out}}^q(z; E_J R') &= -\frac{\alpha_s C_F}{2\pi} \left[ \frac{1}{\varepsilon_{\text{UV}}^2} + \frac{1}{\varepsilon_{\text{UV}}} \left( \frac{3}{2} + \ln \frac{\mu^2}{p_J^{+2} t^2} \right) \right. \\
&\quad \left. + \frac{3}{2} \ln \frac{\mu^2}{p_J^{+2} t^2} + \frac{1}{2} \ln^2 \frac{\mu^2}{p_J^{+2} t^2} + \frac{13}{2} - \frac{3\pi^2}{4} \right].
\end{aligned} \tag{3.20}$$

Note that in some sense this result is trivial, since the integration of the standard jet function in Eq. (3.12) gives the result when there is no restriction of the phase space for the final

state. Because Eq. (3.20) is the same as the integrated jet function for the case  $\theta > R'$ , combining it with  $\mathcal{J}_q(E_J R'; \theta < R')$  in Eq. (3.13) we must have

$$\mathcal{J}_q(E_J R', \theta > R') + \mathcal{J}_q(E_J R', \theta < R') = \int_0^\infty dM^2 J_q(M^2) = 1. \quad (3.21)$$

Thus Eq. (3.20) must have the same result up to a relative minus sign compared with the first order corrections to  $\mathcal{J}_q(E_J R'; \theta < R')$ , obtained from Eqs. (3.13) and (3.14).

The remaining contribution  $[D_{J_q/q}^{\text{out}}(z)]_+$ , is

$$\begin{aligned} [D_{J_q/q}^{\text{out}}(z)]_+ &= [2D_{J_q/q}^{\text{out},(a)}(z) + D_{J_q/q}^{\text{out},(b)}(z)]_+ \\ &= \frac{\alpha_s C_F}{2\pi} \left[ \frac{1+z^2}{1-z} \left( \frac{1}{\varepsilon_{\text{UV}}} + \ln \frac{\mu^2}{p_+^2 t^2} - 2 \ln z(1-z) \right) - (1-z) \right]_+. \end{aligned} \quad (3.22)$$

Combining these results, we arrive at  $D_{J_q/q}^{(1)}$ , i.e., the one loop correction to the quark parton to quark jet fragmentation. Using the identity for the plus distribution,

$$[g(z)h(z)]_+ = [g(z)]_+ h(z) - \delta(1-z) \int_0^1 dy g(y) [h(y) - h(1)], \quad (3.23)$$

we rewrite the renormalized NLO result as

$$\begin{aligned} D_{J_q/q}(z, \mu; ER') &= \delta(1-z) + \frac{\alpha_s C_F}{2\pi} \left\{ \delta(1-z) \left( \frac{3}{2} \ln \frac{\mu^2}{p_+^2 t^2} + \frac{13}{2} - \frac{2\pi^2}{3} \right) - (1-z) \right. \\ &\quad \left. + (1+z^2) \left[ \frac{1}{(1-z)_+} \left( \ln \frac{\mu^2}{p_+^2 t^2} - 2 \ln z \right) - 2 \left( \frac{\ln(1-z)}{1-z} \right)_+ \right] \right\}. \end{aligned} \quad (3.24)$$

We can also compute the contribution for quark parton to gluon jet fragmentation shown in Fig. 11. In this case the gluon in the final state has the momentum  $p_J^+ = zp_+$  and the quark outside the jet has  $(1-z)p_+$ . Therefore the one loop amplitude for  $z \neq 1$  should satisfy the relation  $D_{J_g/q}^{\text{out}}(z) = D_{J_q/q}^{\text{out}}(1-z)$ . Thus the renormalized gluon jet fragmentation function can be written down immediately,

$$D_{J_g/q}(z, \mu; ER') = \frac{\alpha_s C_F}{2\pi} \left[ \frac{1+(1-z)^2}{z} \left( \ln \frac{\mu^2}{p_+^2 t^2} - 2 \ln z(1-z) \right) - z \right]. \quad (3.25)$$

The NLO result for the quark to inclusive FFJ is  $D_{J/q} = D_{J_q/q} + D_{J_g/q}$ , combining Eqs. (3.24) and (3.25). It satisfies the momentum sum rule shown in Eq. (3.4) explicitly. Note that here we expressed the fragmentation functions in terms of  $\ln(\mu^2/p_+^2 t^2)$  rather than

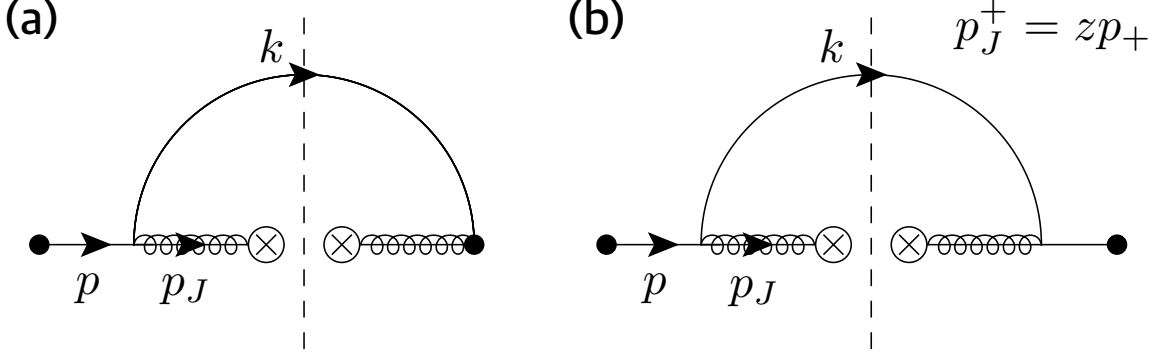


Figure 11: Feynman diagrams for gluon jet splitting contribution at NLO in  $\alpha_s$ . Here the dashed lines represent the unitary cuts. The quark in the final state is outside the jet. Diagram (a) has its Hermitian conjugate contribution.

$\ln(\mu^2/p_J^{+2}t^2)$ . If we rewrite the fragmentation functions with  $\ln(\mu^2/p_J^{+2}t^2)$  using the relation  $p_+ = p_J^+/z$ , these functions cannot satisfy the sum rule in Eq. (3.4) due to additional terms of  $\ln z$ . This fact indicates that the typical scale for the fragmentation function necessary to minimize the large logarithms with small  $R$  is not  $p_J^+t \sim E_J R'$  but  $p_+t \sim ER'$ . For  $z \sim \mathcal{O}(1)$ , the scale choice for FFJ between  $ER$  and  $E_J R$  might not be significant. However the proper choice of the scale can be critical in the small  $z$  limit.

**3.1.2.2 NLO Calculation of Gluon FFJ** As was done for the quark FFJ, we separate the NLO contributions into jet merging and jet splitting contributions. The jet merging contribution is proportional to  $\delta(1-z)$  and includes the virtual contribution. Similarly to Eq. (3.13), the jet merging contribution can be expressed as

$$D_{J/g}^{\text{in}}(z; E_J R') = \delta(1-z) \int_0^{\Lambda^2} dM^2 J_g(M^2; \theta < R') = \delta(1-z) \mathcal{J}_g(E_J R'; \theta < R'), \quad (3.26)$$

where  $\mathcal{J}_g$  is the integrated gluon jet function, which to NLO is given by [46, 47, 48]

$$\begin{aligned} \mathcal{J}_g(E_J R'; \theta < R') = & 1 + \frac{\alpha_s C_A}{2\pi} \left[ \frac{1}{\varepsilon_{\text{UV}}^2} + \frac{1}{\varepsilon_{\text{UV}}} \left( \frac{\beta_0}{2C_A} + \ln \frac{\mu^2}{p_J^{+2}t^2} \right) + \frac{\beta_0}{2C_A} \ln \frac{\mu^2}{p_J^{+2}t^2} \right. \\ & \left. + \frac{1}{2} \ln^2 \frac{\mu^2}{p_J^{+2}t^2} + \frac{67}{9} - \frac{23n_f}{18C_A} - \frac{3\pi^2}{4} \right], \end{aligned} \quad (3.27)$$

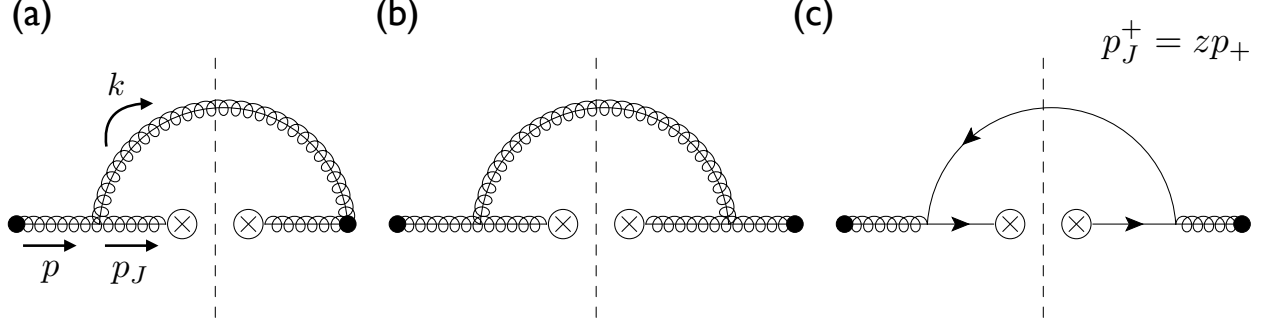


Figure 12: Feynman diagrams for jet splitting contributions to jet fragmentation initiated by gluon parton. Diagram (a) has its Hermitian conjugate contribution. Diagram (a) and (b) represents  $g \rightarrow J_g$  splitting, and Diagram (c) for  $g \rightarrow J_q$  splitting.

where  $C_A = N_c = 3$ , and  $\beta_0 = 11N_c/3 - 2n_f/3$  is the first coefficient of beta function and  $n_f$  is the number of flavors.

In fig. 12, Feynman diagrams for the jet splitting contributions are presented.<sup>3</sup> The contribution of Diagram 12-(a), including the zero-bin subtraction, is

$$\begin{aligned}
D_{J_g/g}^{\text{out,(a)}} = & \frac{\alpha_s C_A}{2\pi} \left\{ \delta(1-z) \left[ -\frac{1}{2\varepsilon_{\text{UV}}^2} - \frac{1}{2\varepsilon_{\text{UV}}} \ln \frac{\mu^2}{p_J^{+2} t^2} - \frac{1}{4} \ln^2 \frac{\mu^2}{p_J^{+2} t^2} + \frac{\pi^2}{24} \right] \right. \\
& + \left( \frac{1}{\varepsilon_{\text{UV}}} + \ln \frac{\mu^2}{p_+^2 t^2} \right) \left[ \frac{z}{(1-z)_+} + \frac{1-z}{z} + \frac{1}{2} \right] \\
& \left. - 2 \left[ \frac{z \ln z}{(1-z)_+} + z \left( \frac{\ln(1-z)}{1-z} \right)_+ + \ln[z(1-z)] \left( \frac{1-z}{z} + \frac{1}{2} \right) \right] \right\}. \quad (3.28)
\end{aligned}$$

The contributions of Diagram 12-(b) is given by

$$D_{J_g/g}^{\text{out,(b)}} = \frac{\alpha_s C_A}{2\pi} \left( \frac{1}{\varepsilon_{\text{UV}}} + \ln \frac{\mu^2}{p_+^2 t^2} - 2 \ln[z(1-z)] \right) (2z(1-z) - 1). \quad (3.29)$$

<sup>3</sup>When we compute Feynman diagrams, we applied the background field method [51], so that no ghost diagrams are involved.

Combining Eqs. (3.26), (3.28), and (3.29), we find NLO result of gluon jet fragmentation function from the gluon,

$$\begin{aligned}
D_{J_g/g}(z, \mu; ER') &= D_{J/g}^{\text{in}} + 2D_{J_g/g}^{\text{out,(a)}} + D_{J_g/g}^{\text{out,(b)}} - \text{UV counter terms} \\
&= \delta(1-z) + \frac{\alpha_s C_A}{2\pi} \left\{ \delta(1-z) \left[ \frac{\beta_0}{2C_A} \ln \frac{\mu^2}{p_+^2 t^2} + \frac{67}{9} - \frac{23n_f}{18C_A} - \frac{2\pi^2}{3} \right] \right. \\
&\quad + 2 \ln \frac{\mu^2}{p_+^2 t^2} \left[ \frac{z}{(1-z)_+} + \frac{1-z}{z} + z(1-z) \right] \\
&\quad \left. - 4 \left[ \frac{z \ln z}{(1-z)_+} + z \left( \frac{\ln(1-z)}{1-z} \right)_+ + \ln[z(1-z)] \left( \frac{1-z}{z} + z(1-z) \right) \right] \right\}. \tag{3.30}
\end{aligned}$$

Diagram 12-(c) contributes to the quark jet fragmentation. The one loop result is given by

$$\begin{aligned}
D_{J_q/g}(z, \mu; ER') &= D_{J_q/g}^{\text{out,(c)}} - \text{UV counter terms} \\
&= \frac{\alpha_s}{2\pi} \left[ \left( \ln \frac{\mu^2}{p_+^2 t^2} - 2 \ln[z(1-z)] \right) \frac{z^2 + (1-z)^2}{2} - z(1-z) \right]. \tag{3.31}
\end{aligned}$$

Note that that Eqs. (3.30) and (3.31) satisfy the momentum conservation sum rule in Eq. (3.4),

$$\int_0^1 dz z \left( [D_{J_g/g}(z) + n_f D_{J_q/g}(z) + n_f D_{J_{\bar{q}/g}(z)}] \right) = \int_0^1 dz z \left( [D_{J_g/g}(z) + 2n_f D_{J_q/g}(z)] \right) = 1. \tag{3.32}$$

**3.1.2.3 Renormalization Scaling Behavior** As can be seen in Eqs. (3.24), (3.25), (3.30), and (3.31), the renormalization group (RG) scaling behavior of the FFJs follows the well-known DGLAP evolution,

$$\frac{d}{d \ln \mu} D_{J_l/k}(x, \mu) = \frac{\alpha_s(\mu)}{\pi} \int_x^1 \frac{dz}{z} P_{lm}(z) D_{J_m/k}(x/z, \mu), \tag{3.33}$$



where the leading splitting kernels are given by

$$P_{qq}(z) = C_F \left[ \frac{3}{2} \delta(1-z) + \frac{1+z^2}{(1-z)_+} \right], \quad (3.34)$$

$$P_{gq}(z) = C_F \left[ \frac{1+(1-z)^2}{z} \right], \quad (3.35)$$

$$P_{qq}(z) = \frac{1}{2} [z^2 + (1-z)^2], \quad (3.36)$$

$$P_{gg}(z) = \frac{\beta_0}{2} \delta(1-z) + 2C_A \left[ \frac{z}{(1-z)_+} + \frac{1-z}{z} + z(1-z) \right]. \quad (3.37)$$

When we compare the higher order result of the FFJ with the fragmentation of a massless parton, the size of the jet,  $ER'$ , suppresses IR sensitivity of the FFJ while the latter has IR divergences. However, both have identical UV behaviors, since the UV divergences arise when the splitting of two particles becomes hard with given large splitting angle.

Comparing to other work, we find that our NLO results for FFJ in Eqs. (3.24), (3.25), (3.30), and (3.31) are the same as “jet functions”,  $j_{k \rightarrow l}$ , in Ref. [52], where the only difference is that the logarithmic terms has been expressed as not  $ER'$  but  $E_J R' = ER'/z$ . This removes the  $\ln z$  term in our expression. However, if we write it this way, we cannot guarantee the momentum sum rule in Eq. (3.4) as we mentioned before. That might give some subtleties for the comparison with other approaches to the estimation of FFJ at higher orders [33, 34].

As noted in the introduction, while completing this work, Ref [41] appeared on the arXiv. The authors have also computed the FFJ at NLO using SCET. The results are the same as ours, but they have the same expression as appearing in Ref. [52]. They claimed that all the virtual diagrams vanish because they are scaleless. However, we believe it is important to carefully separate the UV and IR divergences to obtain a clear picture of the physics. For example, for the case of the jet merging (in-jet) contribution, only when we combine the virtual and real contributions can we obtain an IR finite result.

### 3.1.3 Factorization Theorem for the Fragmentation inside a Jet

To begin, let us consider the scattering cross section with a HFF at a hadron collider:

$$\begin{aligned} \sigma &= \sum_k \int dw dy dp_T \frac{d\sigma_k}{dy dp_T} D_{H/k}(w) \\ &= \sum_k \int dw dy dp_T dp_T^H \frac{d\sigma_k}{dy dp_T} \delta(wp_T^k - p_T^H) D_{H/k}(w), \end{aligned} \quad (3.38)$$

where  $\sigma_k$  is the scattering cross section for the inclusive process with a final parton  $k$ ,  $N_1 N_2 \rightarrow kX$ ,  $p_T$  is the transverse momentum of the parton  $k$  to beam axis,  $y$  is the rapidity of the parton  $k$ , and the rapidity of the hadron can be approximated to be the same as the parton. The differential scattering cross section for the hadron  $H$  is

$$\frac{d\sigma}{dy dp_T^H} = \sum_k \int_{x_H}^1 \frac{dw}{w} \frac{d\sigma_k(y, x_H/w)}{dy dp_T} D_{H/k}(w), \quad (3.39)$$

where  $x_H = p_T^H/Q_T$  and so  $x_H/w = p_T/Q_T$ , with  $Q_T$  being the maximal possible  $p_T$  at a given rapidity.

Next we would like to consider the fragmentation of the hadron inside a jet. In order to do this we factorize the inclusive HFF,

$$D_{H/k}(w) = \sum_l \int_w^1 \frac{dz}{z} B_{J_l/k} \left( \frac{w}{z}; ER' \right) \tilde{D}_{H/J_l}(z; E_J R'), \quad (3.40)$$

where  $J_l$  is the jet with a parton  $l$ , and the momentum fractions are defined as  $z = p_H^+/p_J^+ = p_T^H/p_T^J$  and  $p_J^+/p_+ = p_T^J/p_T = w/z$ .  $B_{J_l/k}$  is the jet splitting kernel from the parton  $k$ , and  $\tilde{D}_{H/J_l}$  is the hadron fragmentation from  $J_l$ .

$\tilde{D}_{H/J_l}$  can be computed by the integration of the fragmenting jet function (FJF) [53, 54],

$$\tilde{D}_{H/J_l}(z; E_J R') = \int_0^{\Lambda^2} dM^2 J_{H/l}(z, M^2), \quad (3.41)$$

where the LO parton level FJF is normalized as  $J_{m/l}(z, M^2) = \delta(1-z)\delta(M^2)\delta_{ml}$ .  $\Lambda^2$  is the maximum jet mass with a given hadron energy fraction  $z$ . For a  $k_T$ -type jet algorithm, it can be expressed as

$$\Lambda_{k_T}^2 = z(1-z)p_J^{+2} \tan^2 \left( \frac{R'}{2} \right). \quad (3.42)$$

The computation of  $\tilde{D}_{H/J_l}(z)$  at NLO was done in Ref. [55, 56]. We also show the NLO calculation in Appendix A.1 separating the UV and IR divergences carefully.

$B_{J_l/k}$  is the jet splitting kernel from the mother parton  $k$ . If we consider the process  $k \rightarrow lm$ , the contribution to  $B_{J_l/k}$  comes from the case where the angle between the partons  $l$  and  $m$  is larger than  $R'$ . Because the convolution of  $B_{J_l/k}$  and  $\tilde{D}_{H/J_l}$  includes all possible fragmentation processes, the result should be the same as the inclusive HFF. However Eq. (3.40) shows that it is possible to describe the whole fragmentation process with a more

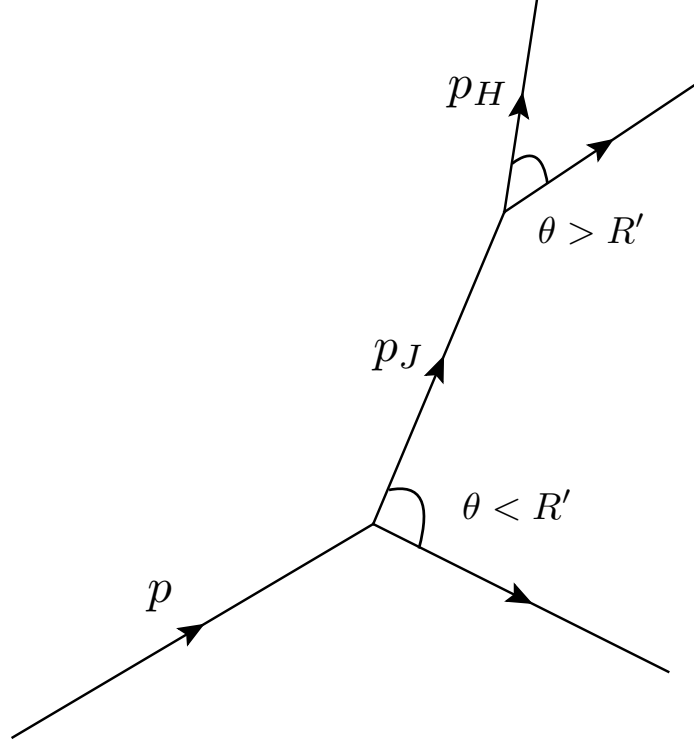


Figure 13: Fragmentation process from the parton ( $p$ ) to the hadron ( $p_H$ ) through the jet ( $p_J$ ).

exclusive observable. The perturbative result of  $B_{J_l/k}$  can be obtained from the matching between  $D_{H/k}$  and  $\tilde{D}_{H/J_l}$ . In Fig. 13 we show the fragmentation process of the hadron through a jet schematically.

We can also consider the direct computation of  $B_{J_l/k}$  based on the calculation of the FFJ in the previous subsection. From the description above,  $B_{J_l/k}$  should be

$$B_{J_l/k}(z, \mu; ER') = \delta(1 - z)\delta_{lk} + D_{J_l/k}^{\text{out}}(z, \mu; ER'), \quad (3.43)$$

where  $D_{J_l/k}^{\text{out}}(z)$  is the jet splitting (out-jet) contribution considered in the FFJ calculation. The typical scale for the jet splitting is  $p_+ \tan(R'/2)$ , which can be approximated as  $ER'$ . Interestingly we find that the perturbative result of the FFJ can be obtained if the higher order result for the jet merging (in-jet) contribution is added to Eq. (3.43). As shown in

Eqs. (3.13) and (3.26), the jet merging contribution can be expressed as  $D_{J/k}^{\text{in}} = \delta(1 - z)\mathcal{J}_k(E_J R')$ . Therefore perturbatively we have the relation <sup>4</sup>

$$D_{J_l/k}(z, \mu; ER') = B_{J_l/k}(z, \mu; ER')\mathcal{J}_l(\mu; E_J R'), \quad (3.44)$$

where the index of the jet parton,  $l$ , is not summed over. Note that the factorized result in Eq. (3.44) has been only confirmed to one loop order. To validate this result beyond NLO, we would need to check the two loop calculation explicitly, which is beyond the scope of this work.

On the right side of Eq. (3.44), having  $\mathcal{J}_l$  rather than  $\mathcal{J}_k$  makes sense beyond NLO accuracy. To see this, consider the case with three final partons at NNLO. If all three particles combine into the jet, the contribution to the FFJ is proportional to  $\delta(1 - z)$ . As seen in Eq. (3.43), the  $\delta_{lk}$  in  $B_{J_l/k}$  guarantees the jet merging contribution is the integrated jet function for the parton  $k$ . However, if we consider the process  $k \rightarrow lm \rightarrow (l_1 l_2)m$  where  $l \rightarrow l_1 l_2$  merged in the jet, this NNLO contribution can be expressed as the multiplication of  $M_{J_l/k}^{\text{out},(1)}$  and  $\mathcal{J}_l^{(1)}$ .<sup>5</sup> Here the superscript (1) denotes the contributions at NLO.

If we apply the momentum conservation sum rule for the hadron to  $\tilde{D}_{H/J_l}(z)$  in Eq. (3.41), we obtain [55]

$$\sum_H \int_0^1 dz z \tilde{D}_{H/J_l}(z) = \int_0^{\Lambda^2} dM^2 \sum_H \int_0^1 dz z J_{H/l}(z, M^2) = \int_0^{\Lambda^2} dM^2 J_l(M^2) = \mathcal{J}_l. \quad (3.45)$$

This also implies the relation of Eq. (3.44). As denoted in Eq. (3.4), the FFJ satisfies the sum rule. Therefore, when applied to Eq. (3.40), the sum rule for the inclusive HFF is guaranteed,

$$\begin{aligned} \sum_H \int_0^1 dw w D_{H/k}(w) &= \sum_l \int_0^1 dx x B_{J_l/k}(x) \sum_H \int_0^1 dz z \tilde{D}_{H/J_l}(z) \\ &= \sum_l \int_0^1 dx x B_{J_l/k}(x) \mathcal{J}_l = \sum_l \int_0^1 dx x D_{J_l/k} = 1. \end{aligned} \quad (3.46)$$

<sup>4</sup>This result has been used for the factorization of the jet mass distribution [57].

<sup>5</sup>We have not considered 3 parton splitting processes at NNLO explicitly. (For the details, see Ref. [58].) It may complicate the factorization in Eq. (3.44).

From Eq.(3.45), we see that the normalization of  $\tilde{D}_{H/J_l}$  is not adequate for a probability. Dividing  $\tilde{D}_{H/J_l}$  by the integrate jet function, we can introduce the HFF inside a jet [56]<sup>6</sup>

$$D_{H/J_l}(z; ER') = \frac{\tilde{D}_{H/J_l}(z, \mu; ER')}{\mathcal{J}_l(\mu; E_J R')} . \quad (3.47)$$

Note that this HFF inside the jet has no renormalization scale dependence because the scale dependence for  $\tilde{D}_{H/J_l}$  is cancelled by  $\mathcal{J}_l$ . (This can be seen by considering the scale dependence in Eq. (3.54) below.) Finally combining Eqs. (3.44) and (3.47) we can rewrite Eq. (3.40) as

$$D_{H/k}(w, \mu) = \sum_l \int_w^1 \frac{dz}{z} D_{J_l/k} \left( \frac{w}{z}, \mu; ER' \right) D_{H/J_l}(z; E_J R'). \quad (3.48)$$

Like a hadron, a jet is also an observable. So it is useful to consider the differential scattering cross section observing the jet and hadron simultaneously. To derive the factorization theorem we combine Eq. (3.39) with Eq. (3.48)

$$\begin{aligned} \frac{d\sigma}{dy dp_T^H} &= \sum_{k,l} \int_{x_H}^1 \frac{dw}{w} \frac{d\sigma_k(y, x_H/w)}{dy dp_T} \\ &\times \int_w^1 \frac{dz}{z} \int dp_T^J \delta(x_J Q_T - p_T^J) D_{J_l/k} \left( \frac{w}{z} \right) D_{H/J_l}(z), \end{aligned} \quad (3.49)$$

where  $x_J = p_T^J/Q_T$ , and we put in the identity  $1 = \int dp_T^J \delta(x_J Q_T - p_T^J)$ . The delta function becomes

$$\delta(x_J Q_T - p_T^J) = \frac{1}{Q_T} \delta \left( x_J - \frac{x_H}{z} \right) = \frac{z^2}{x_H Q_T} \delta \left( z - \frac{x_H}{x_J} \right) . \quad (3.50)$$

Therefore the differential scattering cross section for the jet and the hadron inside the jet can be written as

$$\begin{aligned} \frac{d\sigma}{dy dp_T^J dp_T^H} &= \sum_{k,l} \int_{x_H}^1 \frac{dw}{w} \frac{d\sigma_k(y, p_T/Q_T)}{dy dp_T} \frac{z}{x_H Q_T} D_{J_l/k} \left( \frac{p_T^J}{p_T} \right) D_{H/J_l}(z) \\ &= \sum_{k,l} \int_{x_J}^1 \frac{dx}{x} \frac{d\sigma_k(y, p_T/Q_T = x_J/x)}{dy dp_T} \frac{z}{x_H Q_T} D_{J_l/k}(x) D_{H/J_l}(z). \end{aligned} \quad (3.51)$$

In the second equality we introduced the variable  $x = p_T^J/p_T = w/z$ , hence

$$\int_{x_H}^1 \frac{dw}{w} = \int_{x_J}^1 \frac{dx}{x} . \quad (3.52)$$

---

<sup>6</sup>In Ref. [56], this HFF inside the jet has been called as a jet fragmentation function (JFF).

Finally we have<sup>7</sup>

$$\frac{d\sigma}{dydp_T^J dz} = \sum_{k,l} \int_{x_J}^1 \frac{dx}{x} \frac{d\sigma_k(y, x_J/x)}{dydp_T} D_{J_l/k}(x) D_{H/J_l}(z). \quad (3.53)$$

The factorization theorem in Eq. (3.53) is very useful. For example, instead of the observed hadron, we can consider a subjet inside a fat jet. In this case the factorization theorem becomes

$$\frac{d\sigma}{dydp_T^J dz} = \sum_{k,l} \int_{x_J}^1 \frac{dx}{x} \frac{d\sigma_k(y, x_J/x)}{dydp_T} D_{J_l/k}(x) D_{j/J_l}(z), \quad (3.54)$$

where  $z$  is the momentum fraction of the subjet  $j$  compared to the fat jet  $J$  given by  $z = p_j^+/p_J^+ = p_T^j/p_T^J$  and  $D_{j/J_l}$  is the subjet fragmentation function inside the fat jet. We investigate this more in the following subsection.

### 3.1.4 Subjet Fragmentation inside a Fat jet

For the description of the subjet fragmentation function (sJFF) inside a jet,  $D_{j/J_l}$  in Eq. (3.54), the parton splitting within a fat jet ( $J$ ) with the radius  $R$  only is taken into account. It has a restricted phase space for collinear particle radiations compared to the fully inclusive FFF. As with the HFF inside a jet defined in Eq. (3.47), sFFJ can be written as

$$D_{j/J_l}(z; R'/r') = \frac{\tilde{D}_{j/J_l}(z, \mu; E_J R', R'/r')}{\mathcal{J}_l(\mu; E_J R')}, \quad (3.55)$$

where  $r'$  is the maximal subjet radius. As we will see, the normalized sJFF,  $D_{j/J_l}$ , has no scale dependence except the coupling constant, but depends on the logarithm of  $R'/r'$ .

The naive unnormalized sJFF,  $\tilde{D}_{j/J_l}$ , is described by

$$\begin{aligned} \tilde{D}_{j_k/J_q}(z, \mu) &= \frac{z^{D-3}}{2N_c} \sum_{X_{j-1}, X_{\not{j}}} \text{Tr} \langle 0 | \delta\left(\frac{p_j^+}{z} - \mathcal{P}_+\right) \frac{\not{p}}{2} \Psi_n | j_k(p_j^+, r) X_{\not{j}} \in J(p_J^+, R) \rangle \\ &\times \langle j_k(p_j^+, r) X_{\not{j}} \in J(p_J^+, R) | \bar{\Psi}_n | 0 \rangle, \end{aligned} \quad (3.56)$$

where  $j_k$  represents the subjet with parton  $k$ , and  $r$  is its radius,  $X_{j-1}$  is possible final states within the subjet except the parton  $k$ , and  $X_{\not{j}}$  are the final states not to be included in the

---

<sup>7</sup>In Ref. [52], the similar factorization theorem has been analyzed from the full NLO calculation. We can clearly see the similarity if we express  $D_{H/J_l}$  as Eq. (A.19) when  $\mu \ll E_J R'$ .

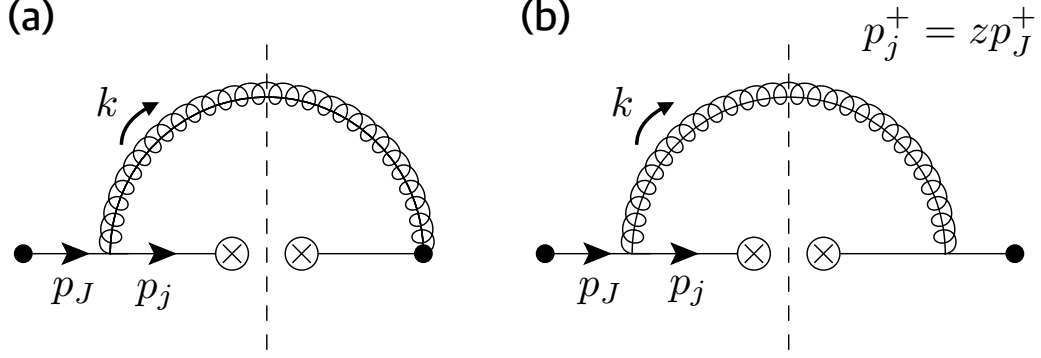


Figure 14: Feynman diagrams of real gluon emissions for the subjet quark fragmentation inside a jet at NLO. Diagram (a) has its Hermitian conjugate contribution.

subjet, but contained in the jet  $J$ . The gluon-initiated sJFF can be expressed similarly in terms of  $\mathcal{B}_n^{\perp a}$  in the adjoint representation.

When we consider the one loop corrections, we will separate the corrections into in-subjet and out-subjet contributions as in Subsection 3.1.2. With the same reasoning as Eq. (3.14), we obtain the in-subjet contribution including the virtual corrections,

$$D_{\text{in}}(z; E_J r') = \delta(1-z) \left\{ 1 + \frac{\alpha_s C_F}{2\pi} \left[ \frac{1}{\varepsilon_{\text{UV}}^2} + \frac{1}{\varepsilon_{\text{UV}}} \left( \frac{3}{2} + \ln \frac{\mu^2}{p_J^{+2} t_r^2} \right) + \frac{3}{2} \ln \frac{\mu^2}{p_J^{+2} t_r^2} + \frac{1}{2} \ln^2 \frac{\mu^2}{p_J^{+2} t_r^2} + \frac{13}{2} - \frac{3\pi^2}{4} \right] \right\}, \quad (3.57)$$

where  $t_r \equiv \tan(r'/2) \sim r'/2$ .

The out-subjet contribution comes from real radiations with  $r' < \theta < R'$ . The naive collinear contribution from the Feynman diagram in Fig. 14-(a) is

$$\begin{aligned} \tilde{D}_{\text{out}}^{(a)}(z) &= \frac{\alpha_s C_F}{2\pi} \frac{(\mu^2 e^\gamma)^\varepsilon}{\Gamma(1-\varepsilon)} \int_{p_J^{+2} t_r^2 z(1-z)}^{p_J^{+2} t_R^2 z(1-z)} \frac{dM^2}{(M^2)^{1+\varepsilon}} z^{1-\varepsilon} (1-z)^{-1-\varepsilon} \\ &= \tilde{I}_{\text{out}}^{(a)} \delta(1-z) + \left[ D_{\text{out}}^{(a)}(z) \right]_+, \end{aligned} \quad (3.58)$$

where  $t_R \equiv \tan(R'/2) \sim R'/2$ , and the tilde represents the result before zero-bin subtractions.

$\tilde{I}_{\text{out}}^{(a)}$  can be extracted by integrating over  $z$ ,

$$\tilde{I}_{\text{out}}^{(a)} = \int_0^1 dz \tilde{D}_{\text{out}}^{(a)}(z) = \frac{\alpha_s C_F}{2\pi} \left[ \left( \frac{1}{2\varepsilon_{\text{IR}}} + 1 \right) \ln \frac{t_r^2}{t_R^2} + \frac{1}{4} \left( \ln^2 \frac{\mu^2}{p_J^{+2} t_R^2} - \ln^2 \frac{\mu^2}{p_J^{+2} t_r^2} \right) \right]. \quad (3.59)$$

Here IR divergence arises as  $z \rightarrow 1$ , which is cancelled by the zero-bin contribution,

$$\begin{aligned} D_{\text{out},0}^{(a)}(z) &= \frac{\alpha_s C_F}{2\pi} \frac{(\mu^2 e^\gamma)^\varepsilon}{\Gamma(1-\varepsilon)} \delta(1-z) \int_0^\infty dk_+ k_+^{-1-\varepsilon} \int_{t_r^2 k_+}^{t_R^2 k_+} k_-^{-1-\varepsilon} \\ &= \frac{\alpha_s C_F}{2\pi} \left[ \frac{1}{2} \left( \frac{1}{\varepsilon_{\text{UV}}} - \frac{1}{\varepsilon_{\text{IR}}} \right) \ln \frac{t_R^2}{t_r^2} \right] \delta(1-z). \end{aligned} \quad (3.60)$$

Hence the IR divergence in Eq. (3.59) is converted to a UV divergence by the zero-bin subtraction.  $[D_{\text{out}}^{(a)}(z)]_+$  is free from IR divergence as  $z \rightarrow 1$  and is given by

$$\left[ D_{\text{out}}^{(a)}(z) \right]_+ = \frac{\alpha_s C_F}{2\pi} \left[ \frac{z}{1-z} \right]_+ \ln \frac{t_R^2}{t_r^2}. \quad (3.61)$$

The out-subjet contribution from diagram Fig. 14-(b) is

$$\begin{aligned} D_{\text{out}}^{(b)}(z) &= \frac{\alpha_s C_F}{2\pi} \frac{(\mu^2 e^\gamma)^\varepsilon}{\Gamma(1-\varepsilon)} (1-\varepsilon) \int_{p_J^{+2} t_r^2 z(1-z)}^{p_J^{+2} t_R^2 z(1-z)} \frac{dM^2}{(M^2)^{1+\varepsilon}} z^{-\varepsilon} (1-z)^{1-\varepsilon} \\ &= I_{\text{out}}^{(b)} \delta(1-z) + \left[ D_{\text{out}}^{(b)}(z) \right]_+, \end{aligned} \quad (3.62)$$

where the terms in the second line are

$$I_{\text{out}}^{(b)} = \int_0^1 dz D_{\text{out}}^{(b)}(z) = \frac{\alpha_s C_F}{2\pi} \left( \frac{1}{2} \ln \frac{t_R^2}{t_r^2} \right), \quad (3.63)$$

$$\left[ D_{\text{out}}^{(b)}(z) \right]_+ = \frac{\alpha_s C_F}{2\pi} (1-z)_+ \ln \frac{t_R^2}{t_r^2}. \quad (3.64)$$

Finally combining Eqs. (3.57), (3.59), (3.60), (3.61), (3.63), and (3.64), we obtain bare NLO result for the naive sJFF:

$$\begin{aligned} \tilde{D}_{j_q/J_q}(z, \mu) &= D_{\text{in}}(z) + 2 \left[ \tilde{D}_{\text{out}}^{(a)}(z) - D_{\text{out},0}^{(a)}(z) \right] + D_{\text{out}}^{(b)}(z) \\ &= \delta(1-z) \left\{ 1 + \frac{\alpha_s C_F}{2\pi} \left[ \frac{1}{\varepsilon_{\text{UV}}^2} + \frac{1}{\varepsilon_{\text{UV}}} \left( \frac{3}{2} + \ln \frac{\mu^2}{p_J^{+2} t_R^2} \right) + \frac{3}{2} \ln \frac{\mu^2}{p_J^{+2} t_R^2} + \frac{1}{2} \ln^2 \frac{\mu^2}{p_J^{+2} t_R^2} \right. \right. \\ &\quad \left. \left. + \frac{13}{2} - \frac{3\pi^2}{4} \right] \right\} + \frac{\alpha_s C_F}{2\pi} \left[ \frac{1+z^2}{1-z} \right]_+ \ln \frac{t_R^2}{t_r^2}. \end{aligned} \quad (3.65)$$



Therefore normalized sJFF can be written as

$$\begin{aligned}
D_{j_q/J_q}(z) &= \frac{\tilde{D}_{j_q/J_q}(z; E_J R', r'/R')}{\mathcal{J}_q(\mu; E_J R')} = \delta(1-z) + \frac{\alpha_s C_F}{2\pi} \left[ \frac{3}{2} \delta(1-z) + \frac{1+z^2}{(1-z)_+} \right] \ln \frac{t_R^2}{t_r^2} \\
&= \delta(1-z) + \frac{\alpha_s}{2\pi} P_{qq}(z) \ln \frac{t_R^2}{t_r^2}.
\end{aligned} \tag{3.66}$$

The gluon subjet fragmentation function from a quark jet can be easily computed. From Eqs. (3.61) and (3.64), exchanging  $z \leftrightarrow 1-z$  and removing ‘+’-distribution we obtain

$$D_{j_g/J_q}(z, \mu) = 2D_{\text{out},g/q}^{(a)}(z) + D_{\text{out},g/q}^{(b)}(z) = \frac{\alpha_s C_F}{2\pi} \frac{1+(1-z)^2}{z} \ln \frac{t_R^2}{t_r^2} = \frac{\alpha_s}{2\pi} P_{gq}(z) \ln \frac{t_R^2}{t_r^2}. \tag{3.67}$$

In a similar manner can compute the sJFFs from the gluon jet. They are given by

$$D_{j_g/J_g}(z) = \delta(1-z) + \frac{\alpha_s}{2\pi} P_{gg}(z) \ln \frac{t_R^2}{t_r^2}, \tag{3.68}$$

$$D_{j_q/J_g}(z) = \frac{\alpha_s}{2\pi} P_{qg}(z) \ln \frac{t_R^2}{t_r^2}. \tag{3.69}$$

If  $t_R \gg t_r$ , the perturbative series expansion fails, and we need to resum the large logarithms of  $t_R/t_r$  to all order in  $\alpha_s$ . To do this, first we integrate out the mode with fluctuations of order  $p^2 \sim p_J^+ t_R^2$ . Then, at the lower scale  $\mu \sim p_J^+ t_r$ , we consider the sJFF setting the upper limit  $p_J^+ t_R \rightarrow \infty$ . Therefore, similar to Eq. (A.19), we obtain the factorization theorem for the subjet fragmentation function

$$D_{j_l/J_k}(z; R'/r') = \int_z^1 \frac{dx}{x} K_{m/k}(z/x, \mu; E_J R') D_{j_l/m}(x, \mu; E_J r'). \tag{3.70}$$

Here  $D_{j_l/m}$  is the standard FFJ for the subjet within the radius  $r$  and the momentum of the mother parton is given by  $p_J$ . The perturbative result is the same as the result in Subsection 3.1.2 with the replacement  $E \rightarrow E_J$  and  $R' \rightarrow r'$ .

The perturbative kernels  $K_{m/k}$  are the matching coefficients between  $D_{j_l/J_k}$  and  $D_{j_l/m}$  and are the result of integrating out the short distance interactions with offshellness  $E_J^2 R'^2$ . They are

$$K_{q/q}(z, \mu) = \delta(1-z) - \frac{\alpha_s}{2\pi} \left\{ P_{qq}(z) \ln \frac{\mu^2}{p_J^{+2} t_R^2} + C_F \left[ \delta(1-z) \left( \frac{13}{2} - \frac{2\pi^2}{3} \right) - (1-z) \right. \right. \\ \left. \left. - 2(1+z^2) \left( \frac{\ln z}{(1-z)_+} + \left( \frac{\ln(1-z)}{1-z} \right)_+ \right) \right] \right\}, \quad (3.71)$$

$$K_{g/q}(z, \mu) = -\frac{\alpha_s}{2\pi} \left[ P_{gq}(z) \left( \ln \frac{\mu^2}{p_J^{+2} t_R^2} - 2 \ln z(1-z) \right) - z C_F \right], \quad (3.72)$$

$$K_{g/g}(z, \mu) = \delta(1-z) - \frac{\alpha_s}{2\pi} \left\{ P_{gg}(z) \ln \frac{\mu^2}{p_J^{+2} t_R^2} + C_A \left[ \delta(1-z) \left( \frac{67}{9} - \frac{23n_f}{18C_A} - \frac{2\pi^2}{3} \right) \right. \right. \\ \left. \left. - 4 \left[ \frac{z \ln z}{(1-z)_+} + z \left( \frac{\ln(1-z)}{1-z} \right)_+ + \ln[z(1-z)] \left( \frac{1-z}{z} + z(1-z) \right) \right] \right] \right\}, \quad (3.73)$$

$$K_{q/g}(z, \mu) = -\frac{\alpha_s}{2\pi} \left[ P_{qg}(z) \left( \ln \frac{\mu^2}{p_J^{+2} t_R^2} - 2 \ln[z(1-z)] \right) - z(1-z) \right]. \quad (3.74)$$

The above results are very interesting. If we replace  $p_J^+$  with the mother parton's momentum,  $p_+$ , we see that the NLO results of  $K_{m/k}$  are the same as NLO corrections to the FFJ with a relative minus sign given in as can be seen from from Eqs. (3.24), (3.25), (3.30), and (3.31). Also, we can see that the sJFF is free from the specific momentum of mother parton, only depending upon the momentum ratio. So, even though there is not much physical meaning, at the computation level we may rewrite Eq. (3.70) as  $D_{j/J}(R'/r') = K(ER') \otimes D_j(Er')$ , with  $\otimes$  is the convolution of the momentum fraction and we show the compatible scale for each function where the compatible scale  $X$  appears in  $\ln(\mu^2/X^2)$  in the NLO calculation.

Based on the results for the factorization theorem in Subsection 3.1.3, let us consider an inclusive scattering cross section for the jet,  $j$  with the radius  $r$  in  $e^+e^-$  annihilation. The scattering cross section is schematically given by

$$\left( \frac{d\sigma}{dE_j} \right)_m = \left( \frac{d\sigma}{dE} \right)_k \otimes [D_j(Er)]_{km} = \left( \frac{d\sigma}{dE} \right)_k \otimes [D_J(ER)]_{kl} \otimes [D_{j/J}(R/r)]_{lm}, \quad (3.75)$$

where the subscripts  $k$ ,  $l$ , and  $m$  denote parton flavors, which are summed for the same indices.  $[D_J]_{kl}$  represents  $D_{J_l/k}$ , and  $[D_{j/J}]_{lm} = D_{j_m/J_l}$ . As discussed below Eq. (3.74)  $[D_{j/J}(R/r)]_{km} = [K(E_J R)]_{kl} \otimes [D_j(E_J r)]_{lm} = [K(ER)]_{kl} \otimes [D_j(Er)]_{lm}$ , where  $[K]_{kl} = K_{l/k}$ . Hence Eq. (3.75) can be written as

$$\begin{aligned} \left(\frac{d\sigma}{dE}\right)_k \otimes [D_j(Er)]_{kn} &= \left(\frac{d\sigma}{dE}\right)_k \otimes [D_J(ER)]_{kl} \otimes [K(ER)]_{lm} \otimes [D_j(Er)]_{mn} \\ &= \left(\frac{d\sigma}{dE}\right)_k \otimes [D_J(ER)]_{kl} \otimes [D_J^{-1}(ER)]_{lm} \otimes [D_j(Er)]_{mn} \quad (3.76) \\ &= \left(\frac{d\sigma}{dE}\right)_m \otimes [D_j(Er)]_{mn}. \end{aligned}$$

This result implies that  $K(ER)$  represents the inverse process of jet fragmentation. This fact demonstrates our observation that the NLO correction to  $K$  putting  $p_+$  instead of  $p_J^+$  is the same as FFJ with the relative minus sign.

Whatever the momentum of the mother parton is, the NLO corrections to FFJ satisfies the sum rule:

$$\sum_l \int_0^1 dz z D_{J_l/k}^{(1)}(z) = 0, \quad (3.77)$$

where again the superscript (1) denotes the NLO correction. Therefore the perturbative kernel  $K_{m/k}$  satisfies the momentum conservation sum rule

$$\sum_m \int_0^1 dz z K_{m/k}(z) = 1. \quad (3.78)$$

### 3.1.5 Conclusions

In this section we introduce the fragmentation function to a jet (FFJ),  $D_{J_k/l}(z, \mu)$ , which describes the fragmentation of a parton  $l$  into a jet with momentum fraction  $z$  with parton  $k$ . This new object naturally appears in factorized rates when considering the jet radius,  $R$ , dependence. To show this, we present a factorization theorem using SCET describing the rate for observing a fragmented hadron and a jet, which is the convolution of the partonic cross section, the FFJ, and the fragmentation of a hadron within a jet as shown in Eq. (3.53).

In order to resum the logarithms of  $R$ , we need the evolution equations for the FFJ. We calculate the NLO corrections for all combinations of quark and gluon initiated to quark

and gluon final state FFJs, and present the results in Eqs. (3.24), (3.25), (3.30), and (3.31). The one loop results of the FFJs satisfy the usual DGLAP evolution equations as seen in Eqs. (3.34-3.37). This allows for the resummation of  $\ln R$  using standard RGE evolutions.

The formalism can be easily generalized to look at other interesting observables. As an example, we show how this formalism can be used to describe a subjet within a fat jet in Eq. (3.54). This allows for the resummation of ratio of the radii of the jets. Using this improved theoretical prediction, we have a better theoretical description of this observable, which may be used to investigate jet substructure as shown in Eqs. (3.66-3.69).

## 3.2 FRAGMENTATION TO A JET IN THE LARGE $Z$ LIMIT

### 3.2.1 Introduction

The fragmentation function (FF) [59] which describes an energetic splitting process for a final state is a very important ingredient in understanding high energy jet and hadron productions. Using the FF we can systematically separate short and long distance interactions related to the productions. For instance an inclusive hadron production for  $e^+e^-$  annihilation can be factorized as

$$\frac{d\sigma(e^+e^- \rightarrow hX)}{dE_h} = \int_{z_h}^1 \frac{dz}{z} \frac{d\sigma_i(z_h/z, \mu)}{dE_i} D_{h/i}(z, \mu), \quad (3.79)$$

where  $i$  denotes the flavor of the produced parton,  $z_h = 2E_h/Q$ , and  $z = E_h/E_i$ . Here  $Q$  is the center of the mass energy of the collision. The partonic scattering cross section  $\sigma_i$  includes hard interactions for  $e^+e^- \rightarrow iX$ . Long distance interactions to describe the fragmenting process from the parton  $i$  to the hadron  $h$  are involved in the FF,  $D_{h/i}(z)$ . The FF is universal in the sense that it is given independently of the hard process and can be applied to other scattering processes. Hence for a long time so many efforts have been made in order to understand the FF deeply (For details we refer to a recent review [60] and the references therein.).

Because we can directly observe a jet using well defined jet algorithms such as ones introduced in Refs. [42, 43, 44, 61, 45], it is also possible to employ the fragmentation function to

a jet (FFJ) instead of a hadron as far as the jet radius  $R$  is enough small [33]. Moreover, once the FFJ for the isolated jet is given, we can systematically investigate its substructures (e.g. hadron and subjet fragmentations [55, 62, 52, 56, 31, 63], and jet mass [57] and transverse momentum [64, 65] distributions) constructing factorization theorems in connection with the fragmenting jet functions [53, 54, 66].

Analytical results of the FFJ have been calculated up to the next-to-leading order in  $\alpha_s$  [52, 41, 31]. Unlike the hadron FF, the FFJ does not have any infrared (IR) divergence because of the finite size of the jet radius  $R$ . However the presence of large logarithms of the small  $R$  does not give a reliable result in perturbation theory and require the resummation to all order in  $\alpha_s$ . As shown in Refs. [33, 52, 41, 31], resumming logarithms of  $R$  is equivalent to scaling down to  $\mu \sim QR$  using Dokshitzer-Gribov-Lipatov-Altarelli-Parisi evolution (DGLAP) equations. This resummed result of the FFJ has been successfully applied to inclusive jet [41, 34] and hadron [63] productions, where the effects of various values of  $R$  have been investigated in detail.

If we observe a highly energetic jet, we might see that most of energetic splitting processes are captured within the jet radius  $R$  since these processes favor narrower angle radiations. This implies that large  $z$  region gives a dominant contribution to the FFJ, where  $z$  is the jet energy fraction over a mother parton energy. Accordingly, in the perturbative result of the FFJ there present large logarithms of  $1 - z$ , which need to be resummed to all order in  $\alpha_s$ . Actually at one loop order there appears a double logarithm such as  $\ln(1 - z)/(1 - z)_+ \sim L^2$ , where  $L$  is a schematic large logarithm. At the accuracy of leading logarithm (LL), it can be resummed as  $\sum_{k=0} C_k(\alpha_s L^2)^k \sim \exp(Lf_0(\alpha_s L))$ , which give a dominant correction to the perturbative expansion of the FFJ.

So for a proper description on the FFJ in the large  $z$  limit we have to systematically handle large logarithms of  $1 - z$  as well as  $R$ . In general if some quantity involves several distinct scales or large logarithms we try to factorize it in order that each factorized part be well described at the scale chosen suitably. Then performing evolutions between largely separated scales, we can properly resum large logarithms. For these purposes we use soft-collinear effective theory (SCET) [16, 17, 19, 20]. SCET provides an appropriate framework for factorization and enable us to resum large logarithms automatically solving

renormalization group (RG) equations for factorized parts.

Near the endpoint where  $z \rightarrow 1$ , we expect the FFJ consists of two dynamics with well separated scales. Since an observed jet has most of energy from a mother parton, the radiation outside the jet should be a soft one with an energy  $\sim E_J(1 - z)$ . Therefore the jet splitting process can be delineated by soft dynamics, while the inside of the jet can be described dominantly by collinear interactions. However, in the effective theory approach wide angle soft interactions are not adequate for explaining the radiation outside the narrow jet because they cannot effectively recognize the jet boundary characterized by small radius  $R$ . Instead, we introduce more refined soft mode, namely the collinear-soft mode [50], which can resolve the narrow jet boundary and describe the out-jet radiations with lower energy consistently. In SCET there have been successful applications of the collinear-soft mode to factorize the cross sections for a narrow jet at low energy scale [37, 38, 39, 40].

In this section we construct a factorization theorem for the FFJ near the endpoint considering collinear and collinear-soft interactions.<sup>8</sup> Then we try to resum large logarithms of  $1 - z$  and  $R$  simultaneously. In Subsection 3.2.2 we discuss the characteristics of large  $z$  physics for the FFJ and factorize it into the collinear and the collinear-soft part. Then, through NLO calculation of each factorized part we confirm our factorized result. In Subsection 3.2.3, based on the factorization, we resum the large logarithms performing RG evolutions of each factorized part. We also discuss large nonglobal logarithms (NGLs) which possibly contribute to the accuracy of NLL. In Subsection 3.2.4 the numerical results of the FFJ to the accuracy of NLL plus NLO in  $\alpha_s$  are shown. In Subsection 3.2.5 finally we conclude and show our outlook.

---

<sup>8</sup> In a strict sense our factorization theorem would hold up to NLO in  $\alpha_s$ . Beyond NLO, large nonglobal logarithms (NGLs) [67, 68] which are relevant to restricted jet phase space might appear and require some modification of our factorization theorem here.

### 3.2.2 The FFJ in the limit $z \rightarrow 1$

First we review the FFJs defined in the previous section:

$$D_{J_k/q}(z, \mu) = \sum_{X_{\notin J}, X_{J-1}} \frac{z^{D-3}}{2N_c} \text{Tr} \langle 0 | \delta \left( \frac{p_J^+}{z} - \mathcal{P}_+ \right) \frac{\vec{n}}{2} \Psi_n | J_k(p_J^+, R) X_{\notin J} \rangle \quad (3.80)$$

$$\begin{aligned} & \times \langle J_k(p_J^+, R) X_{\notin J} | \bar{\Psi}_n | 0 \rangle, \\ D_{J_k/g}(z, \mu) &= \sum_{X_{\notin J}, X_{J-1}} \frac{z^{D-3}}{p_J^+(D-2)(N_c^2-1)} \quad (3.81) \\ & \times \text{Tr} \langle 0 | \delta \left( \frac{P_+}{z} - \mathcal{P}_+ \right) \mathcal{B}_n^{\perp\mu,a} | J_k(p_J^+, R) X_{\notin J} \rangle \langle J_k(p_J^+, R) X_{\notin J} | \mathcal{B}_{n\mu}^{\perp a} | 0 \rangle. \end{aligned}$$

Here  $\Psi_n = W_n^\dagger \xi_n$  and  $\mathcal{B}_n^{\perp\mu,a} = i\bar{n}^\rho g_\perp^{\mu\nu} G_{n,\rho\nu}^b \mathcal{W}_n^{ba} = i\bar{n}^\rho g_\perp^{\mu\nu} \mathcal{W}_n^{\dagger,ba} G_{n,\rho\nu}^b$  are gauge invariant collinear quark and gluon field strength respectively.  $W_n$  ( $\mathcal{W}_n$ ) is a collinear Wilson line in fundamental (adjoint) representation [17, 19]. These collinear fields to describe the jet splitting have a momentum scaling,  $p_n^\mu = (p_+, p_\perp, p_-) = Q(1, \lambda, \lambda^2)$ , where  $\lambda$  is a small parameter comparable to small jet radius  $R$ .  $p_\pm$  are denoted as  $p_+ \equiv \bar{n} \cdot p = p_0 + \hat{\mathbf{n}}_J \cdot \mathbf{p}$  and  $p_- \equiv n \cdot p = p_0 - \hat{\mathbf{n}}_J \cdot \mathbf{p}$ , where  $\hat{\mathbf{n}}_J$  is an unit vector in the jet direction and two lightcone vectors  $n^\mu = (1, \hat{\mathbf{n}}_J)$  and  $\bar{n}^\mu = (1, -\hat{\mathbf{n}}_J)$  have been employed. The expressions for the FFJs in Eqs. (3.80) and (3.81) are valid for the jet frame where the transverse momentum of the observed jet,  $\mathbf{p}_J^\perp$ , is given by zero.

For the specific algorithm to define jet, we will consider inclusive  $k_T$ -type algorithm [42, 43, 44, 45], where merging condition of two light particles are given as

$$\theta < R'. \quad (3.82)$$

Here  $\theta$  is the angle of the two particles, and  $R' = R$  for  $e^+e^-$  collider and  $R' = R/\cosh y$ , where  $y \sim \mathcal{O}(1)$  is the rapidity for the central region.

The definitions of the FFJs in Eq. (3.80) and (3.81) hold for  $z \sim \mathcal{O}(1)$ , but not reliable near the endpoint where  $z$  goes to 1. In the limit  $z \rightarrow 1$ , the observed jet takes most of energy from the mother parton and hence the jet splitting (out-jet) contributions should be described by soft gluon radiations. If  $1-z$  is power counted as  $\mathcal{O}(\eta)$  with  $\eta \ll 1$ , the relevant soft mode would have a momentum scaling such as  $k \sim (k_+, k_\perp, k_-) \sim Q(\eta, \eta, \eta)$ . However, in case of small  $R$ , for the proper resummation of  $\ln R$ , we need a mode to recognize the jet

boundary expressed in terms of  $R$ . This mode would have a lower resolution than the soft mode while  $k_+$  component remains to be power counted as  $\mathcal{O}(\eta)$ . Because the jet merging criterion for the soft gluon radiation is given as [46]

$$\tan^2 \frac{R'}{2} > \frac{k_-}{k_+}, \quad (3.83)$$

the proper mode should allow the hierarchy,  $k_- \sim k_+ \lambda^2 \ll k_+$ , where  $\lambda \sim R$ . Thus this mode should have a scaling,  $k \sim Q\eta(1, \lambda, \lambda^2)$ . From now we will call this mode as the collinear-soft mode.

We can consistently separate the usual soft mode  $\sim Q(\eta, \eta, \eta)$  and the collinear-soft mode as considered in dijet scattering cross section [37, 38]. Also the separation of the collinear-soft mode from collinear fields has been performed in the formulation of SCET<sub>+</sub> [50]. Because the collinear-soft mode can be considered as a subset of the usual soft mode, we have to subtract some overlapped the collinear-soft contribution from the soft contribution in loop calculations similarly with the zero-bin subtraction [49].

If we apply this process to the FFJ with  $z \rightarrow 1$ , we see that the soft contributions can be cancelled by the collinear-soft subtractions. Since the soft mode with a scaling  $(k_+, k_-) \sim Q(\eta, \eta)$  cannot identify the jet boundary in Eq. (3.83), the real soft gluon radiations do not contribute to the in-jet contribution to JFF at all, while the out-jet contribution from real radiations covers the full phase space of  $(k_+, k_-)$ . Thus, without dependence of  $R$ , the total soft contributions should be expressed as a function of  $1 - z$ , namely  $S(1 - z)$ . For the collinear-soft contribution to be subtracted from the soft contribution, we have to apply the same boundary conditions for the soft mode. Hence the real collinear-soft radiations have the only out-jet contributions, which are the same as the case of the soft mode. Therefore the net result of the collinear-soft contributions to be subtracted should be the same as  $S(1 - z)$ , which cancels the soft contribution.

Finally we have a remaining active collinear-soft mode at lower energy scale. When we apply this to the FFJ, we have to keep the jet boundary constraint in Eq. (3.83). As a result the active collinear-soft contributions can be expressed in terms of  $1 - z$  and  $R$  simultaneously. As we will see, the one loop collinear-soft contributions involve the double logarithms of  $\ln \mu / ((1 - z)E_J R')$ . This fact indicates that the collinear-soft interactions



are responsible for large logarithms of  $1 - z$  and its resummation would give a dominant contribution to the FFJ near the endpoint.

**3.2.2.1 Factorization of the FFJ when  $z \rightarrow 1$**  With the reasoning in the above, we can systematically extend the FFJs to the endpoint region including the collinear-soft interactions. We first decouple the soft mode  $\sim Q(\eta, \eta, \eta)$  from the collinear mode  $\sim Q(1, R, R^2)$ . Then we introduce the collinear-soft mode  $\sim Q\eta(1, R, R^2)$  in the collinear sector classifying collinear and collinear-soft gluons such as  $A_n^\mu \rightarrow A_n^\mu + A_{n,cs}^\mu$ . Accordingly the covariant derivative in the collinear sector can be decomposed as  $iD^\mu = iD_c^\mu + iD_{cs}^\mu = \mathcal{P}^\mu + gA_n^\mu + i\partial^\mu + gA_{n,cs}^\mu$ , where  $\mathcal{P}^\mu$  ( $i\partial^\mu$ ) returns collinear (soft-collinear) momentum. In this decomposition, the commutation relations,  $[\mathcal{P}^\mu, A_{n,cs}^\nu] = [\partial^\mu, A_n^\nu] = 0$ , holds. For the factorization of the FFJ, our strategy is simple: After the decomposition into the collinear and the collinear-soft modes, we first integrate out collinear interactions with  $p_c^2 \sim Q^2 R^2$  inside a jet. As we will see, this gives an integrated jet function inside a jet. Then at the lower scale  $\mu_{cs} \sim Q\eta R$  we will consider the collinear-soft interactions for the jet splitting.

As performed in Ref. [50], at the low energy we can additionally introduce so called ‘ultra-collinear’ modes after integrating out collinear interactions with offshellness  $p_c^2 \sim Q^2 R^2$ . These modes have the same order of energy as the collinear mode, but their fluctuations are much smaller than  $Q^2 R^2$ . Then at the low energy scale an external collinear field  $\phi(= \xi, A)_n$  would be matched onto the ultra-collinear fields such as  $\phi_n = \phi_{n_1} + \phi_{n_2} + \dots$ , where the lightcone vectors  $n_{i=1,2,\dots}$  might reside inside the jet with radius  $R$ . Note that collinear interactions between different ultra-collinear modes are forbidden since we already integrate out the large collinear fluctuations  $\sim Q^2 R^2$ . Moreover, as these ultra-collinear modes do very thin collinear interactions, they cannot resolve the jet boundary. Therefore their interactions do not contribute to the FFJs at least to NLO in  $\alpha_s$ . So for simplicity we will not consider ultra-collinear interactions in the FFJ. However, if we have a more refined jet observable to be identified by these mode we have to include them.

Adding the collinear-soft mode, for example, the FFJ with a quark initiation can be

more generically expressed as

$$D_{J_k/q}(z, \mu) = \sum_{X_{\notin J}, X_{J-1}} \frac{z^{D-3}}{2N_c} \text{Tr} \langle 0 | \delta \left( \frac{p_J^+}{z} - \bar{n} \cdot iD \right) \frac{\bar{\eta}}{2} \xi_n | J_k(p_J^+, R) X_{\notin J} \rangle \langle J_k(p_J^+, R) X_{\notin J} | \bar{\xi}_n | 0 \rangle. \quad (3.84)$$

When it compared to Eq. (3.80),  $W_n \delta(p_J^+/z - \mathcal{P}_+) W_n^\dagger = \delta(p_J^+/z - \bar{n} \cdot iD_c)$  has been replaced with  $\delta(p_J^+/z - \bar{n} \cdot iD)$  in Eq. (3.84).

In order to satisfy gauge invariances at each order in  $\lambda \sim \mathcal{O}(R)$  and  $\eta$ , taking the similar procedure considered in Ref. [69] we redefine the collinear gluon field such as

$$A_n^\mu = \hat{A}_n^\mu + \hat{W}_n [iD_{cs}^\mu, \hat{W}_n^\dagger], \quad (3.85)$$

where  $\hat{A}_n$  are newly defined collinear gluon fields and  $\hat{W}_n$  is the collinear Wilson line expressed in terms of  $\hat{A}_n$ . As a consequence the covariant derivative in Eq. (3.84) can be rewritten as

$$iD^\mu = iD_c^\mu + W_n iD_{cs}^\mu W_n^\dagger, \quad (3.86)$$

where collinear fields in the right side are the redefined fields and we removed the hat for simplicity. Employing Eq. (3.86), the delta function in Eq. (3.84) can be rewritten as

$$\delta \left( \frac{p_J^+}{z} - \bar{n} \cdot iD \right) = W_n \delta \left( \frac{p_J^+}{z} - \mathcal{P}_+ - \bar{n} \cdot iD_{cs} \right) W_n^\dagger. \quad (3.87)$$

Similarly with decoupling leading ultrasoft interactions from collinear fields [19], we can remove collinear-soft interactions through the term  $gn \cdot A_{cs}$  in the Lagrangian of the collinear sector. For this, collinear quark and gluon fields can be additionally redefined such that

$$\xi_n \rightarrow Y_n^{cs} \xi_n, \quad A_n^\mu \rightarrow Y_n^{cs} A_n^\mu Y_n^{cs\dagger}, \quad (3.88)$$

where  $Y_n^{cs}$  is the collinear-soft Wilson line to satisfy  $n \cdot iD_{cs} Y_n = Y_n^{cs} n \cdot i\partial$ , and has a similar form with usual soft Wilson lines [19, 70] such as

$$Y_n^{cs}(x) = \text{P} \exp \left[ ig \int_x^\infty ds n \cdot A_{cs}(sn) \right]. \quad (3.89)$$

Using Eqs. (3.86) and (3.88) we can rewrite Eq. (3.84) as

$$D_{J_k/q}(z, \mu) = \sum_{X_{\notin J}, X_{J-1}} \frac{z^{D-3}}{2N_c} \text{Tr} \langle 0 | \delta \left( \frac{p_J^+}{z} - \mathcal{P}_+ - i\partial_+ \right) \frac{\vec{n}}{2} Y_{\bar{n}}^{sc\dagger} Y_n^{sc} W_n^\dagger \xi_n | J_k(p_J^+, R) X_{\notin J} \rangle \\ \times \langle J_k(p_J^+, R) X_{\notin J} | \bar{\xi}_n W_n Y_n^{sc\dagger} Y_{\bar{n}}^{sc} | 0 \rangle, \quad (3.90)$$

where we used the relation  $\bar{n} \cdot iD_{cs} = Y_{\bar{n}}^{cs} i\partial_+ Y_{\bar{n}}^{cs\dagger}$  and  $Y_{\bar{n}}^{cs}$  has the same form as Eq. (3.89) with replacement of  $n \rightarrow \bar{n}$ . We also used the crossing symmetry such as  $\phi \cdots |X_\phi\rangle = \langle X_\phi| \cdots \phi$ , where  $\phi = W_n, Y_{\bar{n}}^{cs}$ . The FFJ in Eq. (3.90) can describe the both regions of ordinary  $z$  and  $z \rightarrow 1$ . If  $z$  is ordinary, i.e.,  $z \sim \mathcal{O}(1)$  and not too close to 1, we can suppress  $i\partial_+$  in the argument of the delta function because  $p_J^+/z - \mathcal{P}_+ \sim \mathcal{O}(Q)$  is power counted much larger than  $i\partial_+ \sim \mathcal{O}(Q\eta)$ . Thus the collinear-soft Wilson lines are cancelled by unitarity and we recover the form in Eq. (3.80). However, in the region  $z \rightarrow 1$ ,  $p_J^+/z - \mathcal{P}_+$  becomes the same size as  $i\partial_+$ . Hence we cannot ignore the term  $i\partial_+$  in the delta function, which gives nonzero contributions of collinear-soft interactions.

Since  $\mathcal{P}_+$  returns collinear (label) momentum in Eq. (3.90),  $\mathcal{P}_+$  can be fixed as  $p_J^+$  near the endpoint. It also means that collinear interactions are relevant to only jet merging (in-jet) contribution to the FFJ. Therefore the FFJ in the limit  $z \rightarrow 1$  can be expressed as

9

$$D_{J_q/q}(z \rightarrow 1, \mu) = \sum_{X_{\notin J}, X_{J-1}} \frac{z^{D-3}}{2N_c} \text{Tr} \langle 0 | Y_{\bar{n}}^{cs\dagger} Y_n^{cs} \frac{\vec{n}}{2} W_n^\dagger \xi_n | J_q(p_J^+, R) X_{\notin J} \rangle \\ \times \langle J_q(p_J^+, R) X_{\notin J} | \bar{\xi}_n W_n \delta \left( \frac{p_J^+}{z} - \mathcal{P}_+^\dagger + i\partial_+ \right) Y_n^{cs\dagger} Y_{\bar{n}}^{cs} | 0 \rangle \\ = \sum_{X_c \in J} \frac{1}{2N_c} \text{Tr} \langle 0 | \frac{\vec{n}}{2} W_n^\dagger \xi_n | q X_c \in J \rangle \langle q X_c \in J | \bar{\xi}_n W_n | 0 \rangle \cdot \sum_{X_{cs}} \frac{1}{N_c} \text{Tr} \langle 0 | Y_{\bar{n}}^{cs\dagger} Y_n^{cs} | X_{cs} \rangle \\ \times \langle X_{cs} | \delta((1-z)p_J^+ + \Theta(\theta - R') i\partial_+) Y_n^{cs\dagger} Y_{\bar{n}}^{cs} | 0 \rangle, \quad (3.91)$$

where  $\Theta$  is the step function and we reorganized the final states in the first equality as the collinear states ( $q X_c$ ) in the jet and collinear-soft states  $X_{cs}$  in order to factorize collinear

---

<sup>9</sup>Note that the splitting  $q \rightarrow J_g$  in the limit  $z \rightarrow 1$  is power suppressed by  $\mathcal{O}(1-z)$  compared to the splitting  $q \rightarrow J_q$ . For  $q \rightarrow J_g$ , the splitted parton away from the observed jet is the soft-collinear quark, which gives a power suppression of  $\mathcal{O}(\eta)$  compared to the soft-collinear gluon radiation. Also in case of the gluon splitting,  $g \rightarrow J_g$  is dominant for the same reason.

and collinear-soft interactions. In the second equality we fixed the collinear label momentum  $\mathcal{P}^\dagger$  as  $p_J^+$ . And then we put the jet splitting constraint in front of  $i\partial_+$  because only out-jet collinear-soft radiations gives a nonzero contribution for the region  $z < 1$ . From Eq. (3.83), the jet splitting constraint  $\Theta(\theta - R')$  is equivalent to  $\tan^2 R'/2 < k_-/k_+$ , where  $k$  is the collinear-soft momentum.

Eq. (3.91) shows that the quark FFJ in the limit  $z \rightarrow 1$  is factorized as

$$D_{J_q/q}(z \rightarrow 1, \mu; E_J R', (1-z)E_J R') = \mathcal{J}_q(\mu; E_J R', \theta < R') \cdot S_q(z, \mu; (1-z)E_J R'), \quad (3.92)$$

where  $\mathcal{J}_q$  is the integrated jet function for the in-jet contribution and defined as

$$\mathcal{J}_q(\mu; E_J R', \theta < R') = \sum_{X_c \in J} \frac{1}{2N_c p_J^+} \text{Tr} \langle 0 | \frac{\not{p}}{2} W_n^\dagger \xi_n | q X_c \in J(E_J, R') \rangle \langle q X_c \in J | \bar{\xi}_n W_n | 0 \rangle. \quad (3.93)$$

$S_q$  is the dimensionless collinear-soft function. When we rewrite  $S_q = p_J^+ \tilde{S}_q$ , the dimensionful collinear-soft function  $\tilde{S}_q$  can be expressed as

$$\tilde{S}_q(\ell_+, \mu; \ell_+ t) = \sum_{X_{cs}} \frac{1}{N_c} \text{Tr} \langle 0 | Y_{\bar{n}}^{cs\dagger} Y_n^{cs} | X_{cs} \rangle \langle X_{cs} | \delta(\ell_+ + \Theta(\theta - R') i\partial_+) Y_n^{cs\dagger} Y_{\bar{n}}^{cs} | 0 \rangle, \quad (3.94)$$

where  $t \equiv \tan R'/2$ , and  $\ell_+ t$  is the scale to minimize large logarithms in the higher order corrections as we will see later.

Using the adjoint representation and taking a similar procedure with the quark case, we also have the factorization formula for the gluon FFJ such that

$$D_{J_g/g}(z \rightarrow 1, \mu) = \mathcal{J}_g(\mu; E_J R', \theta < R') \cdot S_g(z, \mu; (1-z)E_J R'), \quad (3.95)$$

where  $\mathcal{J}_g$  is the integrated jet function for gluon. And  $S_g$  is the collinear-soft function defined similarly with Eq.(3.94), where the Wilson lines in the adjoint representation can be used instead of  $Y_{n,\bar{n}}^{cs}$ .

**3.2.2.2 NLO calculation of the FFJ near the endpoint** The integrated jet functions shown in Eqs. (3.92) and (3.95) has been explicitly computed at NLO [47, 46, 48] and partially computed at NNLO [38, 39]. The NLO results with the constraint of Eq. (3.82) read

$$\mathcal{J}_q(\mu; E_J R', \theta < R') = 1 + \frac{\alpha_s C_F}{2\pi} \left[ \frac{3}{2} \ln \frac{\mu^2}{p_J^{+2} t^2} + \frac{1}{2} \ln^2 \frac{\mu^2}{p_J^{+2} t^2} + \frac{13}{2} - \frac{3\pi^2}{4} \right], \quad (3.96)$$

$$\begin{aligned} \mathcal{J}_q(\mu; E_J R', \theta < R') = 1 + \frac{\alpha_s C_A}{2\pi} & \left[ \frac{1}{\varepsilon_{\text{UV}}^2} + \frac{1}{\varepsilon_{\text{UV}}} \left( \frac{\beta_0}{2C_A} + \ln \frac{\mu^2}{p_J^{+2} t^2} \right) + \frac{\beta_0}{2C_A} \ln \frac{\mu^2}{p_J^{+2} t^2} \right. \\ & \left. + \frac{1}{2} \ln^2 \frac{\mu^2}{p_J^{+2} t^2} + \frac{67}{9} - \frac{23n_f}{18C_A} - \frac{3\pi^2}{4} \right], \end{aligned} \quad (3.97)$$

where  $p_J^+ t \sim E_J R'$ ,  $\beta_0 = 11N_c/3 - 2n_f/3$ ,  $C_A = N_c = 3$ ,  $n_f$  is the number of flavors.

For NLO computation of the collinear-soft function in Eq. (3.94) we consider virtual and real gluon contributions respectively. When we separate ultraviolet (UV) and infrared (IR) divergences carefully, the virtual contributions are given by

$$M_V^S = -\frac{\alpha_s C_F}{\pi} \left( \frac{1}{\varepsilon_{\text{UV}}} - \frac{1}{\varepsilon_{\text{IR}}} \right)^2 \delta(\ell_+). \quad (3.98)$$

The real contributions at one loop can be written as

$$\begin{aligned} M_R^S &= \frac{\alpha_s C_F}{\pi} \frac{(\mu^2 e^{\gamma_E})^\epsilon}{\Gamma(1-\epsilon)} \int_0^\infty dk_+ dk_- (k_+ k_-)^{-1-\epsilon} \left[ \delta(\ell_+ - k_+) \Theta(k_- - t^2 k_+) \right. \\ & \quad \left. + \delta(\ell_+) \Theta(t^2 k_+ - k_-) \right] \equiv M_{R1}^S + M_{R2}^S, \end{aligned} \quad (3.99)$$

where  $k$  is the momentum of the outgoing collinear-soft gluon and  $M_{R1}^S$  ( $M_{R2}^S$ ) indicates the contribution from the first (second) term in the bracket.

In Fig. 15 we show the possible phase space for the emitted soft-collinear gluon after the integration on  $\mathbf{k}_\perp$ .  $M_{R2}^S$  covers lower region from the jet border line ( $k_- = t^2 k_+$ ). Hence the result is obtained as

$$\begin{aligned} M_{R2}^S &= \frac{\alpha_s C_F}{\pi} \frac{(\mu^2 e^{\gamma_E})^\epsilon}{\Gamma(1-\epsilon)} \delta(\ell_+) \int_0^\infty dk_+ \int_0^{t^2 k_+} dk_- (k_+ k_-)^{-1-\epsilon} \\ &= \frac{\alpha_s C_F}{2\pi} \left[ \left( \frac{1}{\varepsilon_{\text{UV}}} - \frac{1}{\varepsilon_{\text{IR}}} \right)^2 - \left( \frac{1}{\varepsilon_{\text{UV}}} - \frac{1}{\varepsilon_{\text{IR}}} \right) \ln t^2 \right] \delta(\ell_+). \end{aligned} \quad (3.100)$$

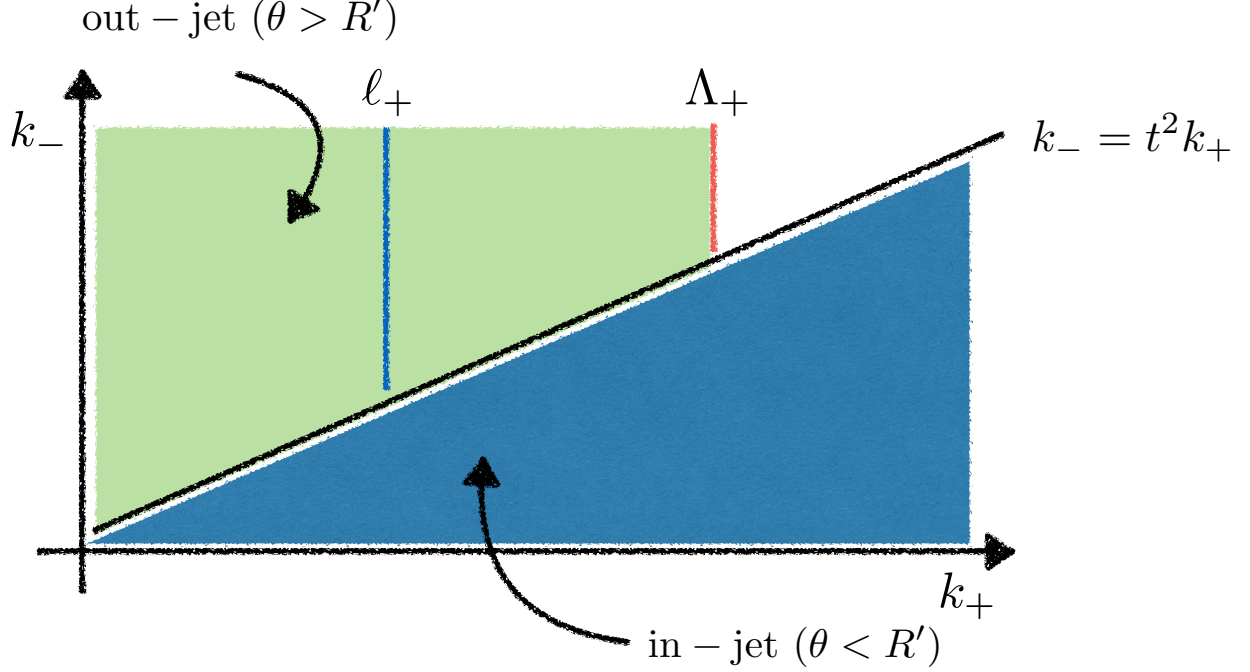


Figure 15: Phase space for the real gluon emission in the soft function. In  $(k_+, k_-)$  plane the region above the border line  $k_- = t^2 k_+$  gives out-jet contribution and the region in the below gives in-jet contribution.  $\Lambda_+$  is the maximum value for the distribution of  $\ell_+$  and can be chosen arbitrarily.

For  $M_{R1}^S$ ,  $k_+$  has been fixed as  $\ell_+$ , and the possible phase space has been denoted as a blue line in the upper plane in Fig 15. However we need to extract IR divergences as  $\ell_+ \rightarrow 0$ . In order to do that, we introduce so called  $\Lambda_+$ -distribution, which is defined as

$$\int_0^L d\ell_+ [g(\ell_+)]_{\Lambda_+} f(\ell_+) = \int_0^L d\ell_+ g(\ell_+) f(\ell_+) - \int_0^{\Lambda_+} d\ell_+ g(\ell_+) f(0), \quad (3.101)$$

where  $f(\ell_+)$  is an arbitrary smooth function at  $\ell_+ = 0$ .  $\Lambda_+$  is an arbitrary upper limit for  $\Lambda_+$ -distribution and power counted to have the same size as  $\ell_+$ . Using the distribution we can write  $M_{R1}^S$  such that

$$\begin{aligned} M_{R1}^S &= \frac{\alpha_s C_F}{\pi} \frac{(\mu^2 e^{\gamma_E})^\epsilon}{\Gamma(1-\epsilon)} \ell_+^{-1-\epsilon} \int_{t^2 \ell_+}^\infty dk_- k_-^{-1-\epsilon} \\ &= \delta(\ell_+) I_{R1}(\Lambda_+, t) + \frac{\alpha_s C_F}{\pi} \frac{(\mu^2 e^{\gamma_E})^\epsilon}{\Gamma(1-\epsilon)} \left[ \ell_+^{-1-\epsilon} \int_{t^2 \ell_+}^\infty k_-^{-1-\epsilon} \right]_{\Lambda_+}, \end{aligned} \quad (3.102)$$

where the integration region for  $I_{R1}$  is corresponding to a green region in Fig. 15. It is computed as

$$\begin{aligned}
I_{R1} &= \frac{\alpha_s C_F}{\pi} \frac{(\mu^2 e^{\gamma_E})^\epsilon}{\Gamma(1-\epsilon)} \left[ \int_0^\infty dk_+ \int_{t^2 k_+}^\infty dk_- (k_+ k_-)^{-1-\epsilon} - \int_{\Lambda_+}^\infty dk_+ \int_{t^2 k_+}^\infty dk_- (k_+ k_-)^{-1-\epsilon} \right] \\
&= \frac{\alpha_s C_F}{2\pi} \left[ \left( \frac{1}{\epsilon_{UV}} - \frac{1}{\epsilon_{IR}} \right)^2 + \left( \frac{1}{\epsilon_{UV}} - \frac{1}{\epsilon_{IR}} \right) \ln t^2 \right. \\
&\quad \left. - \left( \frac{1}{\epsilon_{UV}^2} + \frac{1}{\epsilon_{UV}} \ln \frac{\mu^2}{\Lambda_+^2 t^2} + \frac{1}{2} \ln^2 \frac{\mu^2}{\Lambda_+^2 t^2} - \frac{\pi^2}{12} \right) \right].
\end{aligned} \tag{3.103}$$

The second term in Eq. (3.102) is given by

$$\frac{\alpha_s C_F}{\pi} \frac{(\mu^2 e^{\gamma_E})^\epsilon}{\Gamma(1-\epsilon)} \left[ \ell_+^{-1-\epsilon} \int_{t^2 \ell_+}^\infty k_-^{-1-\epsilon} \right]_{\Lambda_+} = \frac{\alpha_s C_F}{\pi} \left[ \frac{1}{\ell_+} \left( \frac{1}{\epsilon_{UV}} + \ln \frac{\mu^2}{\ell_+^2 t^2} \right) \right]_{\Lambda_+}. \tag{3.104}$$

Finally combining Eqs. (3.98), (3.100), (3.103) and (3.104) we obtain the bare one loop result of  $\tilde{S}_q$  such as

$$\begin{aligned}
M_S &= M_V^S + M_{R1}^S + M_{R2}^S \\
&= \frac{\alpha_s C_F}{\pi} \left\{ \delta(\ell_+) \left( -\frac{1}{2\epsilon_{UV}^2} - \frac{1}{2\epsilon_{UV}} \ln \frac{\mu^2}{\Lambda_+^2 t^2} - \frac{1}{4} \ln^2 \frac{\mu^2}{\Lambda_+^2 t^2} + \frac{\pi^2}{24} \right) \right. \\
&\quad \left. + \left[ \frac{1}{\ell_+} \left( \frac{1}{\epsilon_{UV}} + \ln \frac{\mu^2}{\ell_+^2 t^2} \right) \right]_{\Lambda_+} \right\}.
\end{aligned} \tag{3.105}$$

The one loop result of the collinear-soft function for gluon FFJ is the same if we replace  $C_F$  with  $C_A = N_c$  in Eq. (3.105).

Since the dimensionless soft-collinear function,  $S_{k=q,g}(z) = p_J^+ \tilde{S}_k(\ell_+)$ , is the function of  $z$ , we need to express  $\Lambda_+$ -distribution as the standard plus distribution of  $z$ . From Eq. (3.101) we obtain the relation

$$[\tilde{g}(\ell_+)]_{\Lambda_+} = \frac{1}{p_J^+} [g(z)]_+ + \frac{1}{p_J^+} \delta(1-z) \int_0^b dz' g(z'), \tag{3.106}$$

where  $\ell_+ = p_J^+(1-z)$  and  $g(z) = p_J^+ \tilde{g}(\ell_+)$ . Also in  $\Lambda_+$ -distribution  $\Lambda_+$  has been replaced with  $p_J^+(1-b)$ , where  $b$  is a dimensionless parameter close to 1.

After all the dimensionless soft-collinear functions at NLO can be written as follows:

$$\begin{aligned}
S_{k=q,g}(z, \mu; (1-z)E_J R') &= \delta(1-z) + \frac{\alpha_s C_k}{2\pi} \left\{ \delta(1-z) \left( -\frac{1}{2} \ln^2 \frac{\mu^2}{p_J^{+2} t^2} + \frac{\pi^2}{12} \right) \right. \\
&\quad \left. + 2 \left[ \frac{1}{(1-z)} \left( \ln \frac{\mu^2}{p_J^{+2} t^2} - 2 \ln(1-z) \right) \right]_+ \right\}, \quad (3.107)
\end{aligned}$$

where  $C_q = C_F$  and  $C_g = C_A$ . As seen in Eqs. (3.105) and (3.107), the scale for the soft-collinear function to minimize the large logarithms is given as  $(1-z)E_J R'$ . So in the limit  $z \rightarrow 1$  the scale evolution to  $(1-z)E_J R'$  with double logarithms is inevitable for the precise estimation of the FFJ. In Eqs. (3.92) and (3.95) we have shown the factorization theorem near the endpoint. Combining Eqs. (3.96), (3.97) and (3.107) we can easily check that the fixed NLO result of Eqs. (3.92) and (3.95) can recover the NLO results of FFJs for the full range [31, 52, 41] when we take the limit  $z \rightarrow 1$ .

### 3.2.3 Renormalization Group Evolution and Resummation of Large Logarithms

**3.2.3.1 RG evolution from the factorization of the FFJ** Based on the factorized results in Eqs. (3.92) and (3.95), we can systematically resum the large logarithms such as  $\ln R$  and  $\ln(1-z)$  in the FFJ through renormalization group (RG) evolutions of the integrated jet function  $\mathcal{J}_k$  and the collinear-soft functions  $S_k$ . The FFJ in the limit  $z \rightarrow 1$  can be factorized at an arbitrary factorization scale  $\mu_f$ . Then  $\mathcal{J}_k$  can be evolved from  $\mu_f$  to collinear scale  $\mu_c \sim E_J R'$ , where the large logarithms at the higher order in  $\alpha_s$  are minimized and the perturbative expansion is safely convergent. Also we can evolve  $S_k$  from  $\mu_f$  to  $\mu_{cs} \sim (1-z)E_J R'$  to minimize the large logarithms at  $\mu_{cs}$ . Because the fixed order results of the integrated jet function at  $\mu_c$  and the collinear-soft function at  $\mu_{cs}$  do not involve large logarithms any more, RG evolutions from  $\mu_f$  to  $\mu_c$  and  $\mu_{cs}$  can automatically include resumming whole large logarithms, and the final result can be expressed as an exponentiation form of the large logarithms.



Anomalous dimensions of the integrated jet functions and the collinear-soft functions are described as

$$\frac{d}{d \ln \mu} \mathcal{J}_k(\mu) = \gamma_{c,k}(\mu) \mathcal{J}_k(\mu), \quad (3.108)$$

$$\frac{d}{d \ln \mu} S_k(x, \mu)(\mu) = \int_x^1 \frac{dz}{z} \gamma_{cs,k}(z, \mu) S_k(x/z, \mu), \quad (3.109)$$

where  $k = q, g$ . From Eqs. (3.96), (3.97), and (3.107), the anomalous dimensions at one loop are given as

$$\gamma_{c,q}^{(0)} = \frac{\alpha_s C_F}{2\pi} \left( 2 \ln \frac{\mu^2}{E_J^2 R'^2} + 3 \right), \quad \gamma_{c,g}^{(0)} = \frac{\alpha_s C_A}{2\pi} \left( 2 \ln \frac{\mu^2}{E_J^2 R'^2} + \frac{\beta_0}{C_A} \right), \quad (3.110)$$

$$\gamma_{cs,k}^{(0)}(z) = \frac{\alpha_s C_k}{2\pi} \left( -2 \ln \frac{\mu^2}{E_J^2 R'^2} \delta(1-z) + \frac{4}{(1-z)_+} \right), \quad (3.111)$$

where  $p_J^+ t$  is approximated as  $E_J R'$ . When we sum Eqs. (3.110) and (3.111), the logarithmic terms are cancelled and the well-known DGLAP splitting kernels in the limit  $z \rightarrow 1$  are reproduced such as

$$\frac{\alpha_s}{\pi} P_{kk}^{(0)}(z \rightarrow 1) = \delta(1-z) \gamma_{c,k}^{(0)} + \gamma_{cs,k}^{(0)}(z). \quad (3.112)$$

Logarithmic terms in the leading anomalous dimensions indicate the presence of the cusp anomalous dimension. Beyond leading logarithmic (LL) accuracy the anomalous dimensions can be expressed as

$$\gamma_{c,k} = A_c \Gamma_{C,k}(\alpha_s) \ln \frac{\mu^2}{E_J^2 R'^2} + \hat{\gamma}_{c,k}(\alpha_s), \quad (3.113)$$

$$\gamma_{cs,k}(z) = \delta(1-z) \left[ A_{cs} \Gamma_{C,k}(\alpha_s) \ln \frac{\mu^2}{E_J^2 R'^2} + \hat{\gamma}_{cs,k}(\alpha_s) \right] - \kappa_{cs} A_{cs} \frac{\Gamma_{C,k}(\alpha_s)}{(1-z)_+}, \quad (3.114)$$

where  $\Gamma_{C,k} = \sum_{n=0} \Gamma_{n,k}(\alpha_s/4\pi)^{n+1}$  are the cusp anomalous dimensions obtained from calculations of the light like Wilson loops [71, 72]. The first two coefficients are given by

$$\Gamma_{0,k} = 4C_k, \quad \Gamma_{1,q} = 4C_k \left[ \left( \frac{67}{9} - \frac{\pi^2}{3} \right) C_A - \frac{10}{9} n_f \right]. \quad (3.115)$$

From LO results in Eqs. (3.110) and (3.111) we extract  $\{A_c, A_{cs}, \kappa_{cs}\} = \{1, -1, 2\}$  and the noncusp anomalous dimensions such as  $\hat{\gamma}_{c,q} = 3\alpha_s C_F/(2\pi) + \mathcal{O}(\alpha_s^2)$ ,  $\hat{\gamma}_{c,g} = \alpha_s \beta_0/(2\pi) + \mathcal{O}(\alpha_s^2)$ , and  $\hat{\gamma}_{cs,k} = \mathcal{O}(\alpha_s^2)$ .

Using Eqs. (3.113) and (3.114) we perform RG evolutions of the integrated jet functions and the collinear-soft functions up to next-to-leading logarithmic (NLL) accuracy. For  $\mathcal{J}_k$  the result of RG evolution from  $\mu_f$  to  $\mu$  can be written as

$$\mathcal{J}_k(\mu_f) = \exp \left[ 2A_c S_\Gamma(\mu_f, \mu_c) + A_c \ln \frac{\mu_f^2}{E_J^2 R'^2} a[\Gamma_{C,k}](\mu_f, \mu_c) + a[\hat{\gamma}_{c,k}](\mu_f, \mu_c) \right] \mathcal{J}_k(\mu_c). \quad (3.116)$$

Here  $S_\Gamma$  and  $a[f]$  are expressed as

$$S_\Gamma(\mu_f, \mu_c) = \int_{\alpha_c}^{\alpha_f} \frac{d\alpha_s}{b(\alpha_s)} \Gamma_{C,k}(\alpha_s) \int_{\alpha_f}^{\alpha_s} \frac{d\alpha'_s}{b(\alpha'_s)}, \quad a[f](\mu_f, \mu_c) = \int_{\alpha_c}^{\alpha_f} \frac{d\alpha_s}{b(\alpha_s)} f(\alpha_s), \quad (3.117)$$

where  $\alpha_{f,c} \equiv \alpha_s(\mu_{f,c})$  and  $b(\alpha_s) = d\alpha_s/(d \ln \mu)$  is QCD beta function.

For the evolution of  $S_k$ , following the conventional method introduced in Refs. [73, 74] we obtain

$$\begin{aligned} S_k(z, \mu_f) &= \exp \left[ 2A_{cs} S_\Gamma(\mu_f, \mu_{cs}) + a[\hat{\gamma}_{cs,k}](\mu_f, \mu_{cs}) \right] \left( \frac{\mu_f^2}{E_J^2 R'^2} \right)^{-\eta_S/\kappa_{cs}} \\ &\times \bar{S}_k \left[ \ln \frac{\mu_{cs}^2}{E_J^2 R'^2} - 2\partial_{\eta_S} \right] \frac{e^{-\gamma_E \eta_S}}{\Gamma(\eta_S)} (1-z)^{(-1+\eta_S)}, \end{aligned} \quad (3.118)$$

where  $\eta_S$  is defined as  $\eta_S = -\kappa_{cs} A_{cs} a[\Gamma_{C,k}](\mu_f, \mu_{cs})$  and given by a positive number for  $\mu_f > \mu_{cs}$ .  $\bar{S}_k$  is obtained as

$$\bar{S}_k[L] = 1 + \frac{\alpha_s C_k}{2\pi} \left( -\frac{1}{2} L^2 - \frac{\pi^2}{4} \right) + \mathcal{O}(\alpha_s^2). \quad (3.119)$$

**3.2.3.2 Contribution of nonglobal logarithms** When we extend the factorized result of the FFJ to the two loop or the higher order in  $\alpha_s$ , one important issue is the presence of nonglobal logarithms (NGLs) [67, 68]. Usually the NGLs appears when jet observables only reflect a limited phase space due to a jet algorithm, and arises from multiple gluon radiations near the jet boundary. Especially when there are large energy differences of the radiated gluons between to in-jet and to out-jet, the large NGLs are unavoidable.

In case of the FFJ near the endpoint there are two modes to resolve a jet boundary and to give nonvanishing contributions: the collinear mode with a large energy certainly radiates only inside a jet, and the collinear-soft mode to radiate across a jet boundary gives a nonvanishing result of  $1 - z$  at the lower energy scale. So we properly guess there can exist the large NGLs in the FFJ in the large  $z$  limit.

In order to systematically resum the large NGLs, we expect our factorization theorem be modified because it is designed to resum global Sudakov logarithms. For inclusion of resummation of NGLs in the effective theory approach, from two loop order we might have to consider dressed collinear-soft gluons decoupled from a (ultra-)collinear gluon in a certain direction inside a jet, which might give a new dipole operator other than  $Y_{n,\bar{n}}^{cs}$  at the lower energy. We will not pursue such a refined factorization theorem here, but we mention that some advanced treatments of the NGLs have been recently introduced in Refs. [37, 39, 75, 76, 77, 78, 79].

The NGL in FFJ would take the same form as the endpoint logarithms,  $\ln(1 - z)$ , and it can be inferred from the scale ratio between the typical collinear scale  $\mu_c \sim E_J R'$  and the collinear-soft scale  $\mu_{cs} \sim (1 - z)ER'$ . As seen in the threshold expansion of inclusive jet production [80], leading NGL starts to appear at two loop as  $\alpha_s^2 L^2 \sim \alpha_s^2 (\ln(1 - z)/(1 - z))_+$ , where  $L$  denotes a large logarithm schematically. So at NLL accuracy we have to resum these leading NGLs to all order in  $\alpha_s$  such as  $\sum_{n=2} C_{\text{NG}}^n(\alpha_s L)^n$ .

In case of hemisphere jet mass distribution in  $e^+e^-$  annihilation, the resummed result of leading NGLs has been known in the large  $N_c$  limit [67]. Interestingly the resummed result of leading NGLs for an individual narrow jet is found to have the same form as the case of hemisphere jet mass, and the only difference simply arises from evolution scales to be chosen suitably [81, 82]. Therefore, using the result in Ref. [67] we guess the resummed result of

leading NGLs for the FFJ in the large  $N_c$  limit such as

$$\Delta_{\text{NG}}^k(\mu_c, \mu_{cs}) = \exp\left(-C_A C_k \frac{\pi^2}{3} \left(\frac{1 + (at)^2}{1 + (bt)^c}\right) t^2\right), \quad (3.120)$$

where  $k = q, g$ , and

$$t = \frac{1}{\beta_0} \ln \frac{\alpha_s(\mu_{cs})}{\alpha_s(\mu_c)} \sim -\frac{1}{\beta_0} \ln\left(1 - \frac{\beta_0}{4\pi} \alpha_s(\mu_c) \ln \frac{\mu_c^2}{\mu_{cs}^2}\right). \quad (3.121)$$

The fit parameters from the Monte Carlo implementation of the parton-shower are given as  $a = 0.85C_A$ ,  $b = 0.86C_A$ ,  $c = 1.33$  [67].

Up to NLL accuracy (plus NLO in  $\alpha_s$ ), the resummation factor for NGLs in Eq. (3.120) can be practically multiplied to the resummed results of the FFJ in the previous subsection, where the resummed expressions of  $\mathcal{J}_{k=q,g}$  and  $S_{k=q,g}$  are shown in Eqs. (3.116) and (3.118) respectively. In the next subsection we show various numerical results of the FFJ in the large  $z$  region comparing the results using only DGLAP evolutions and our resummed results of the large logarithms as well as the NGLs.

### 3.2.4 Numerical Results

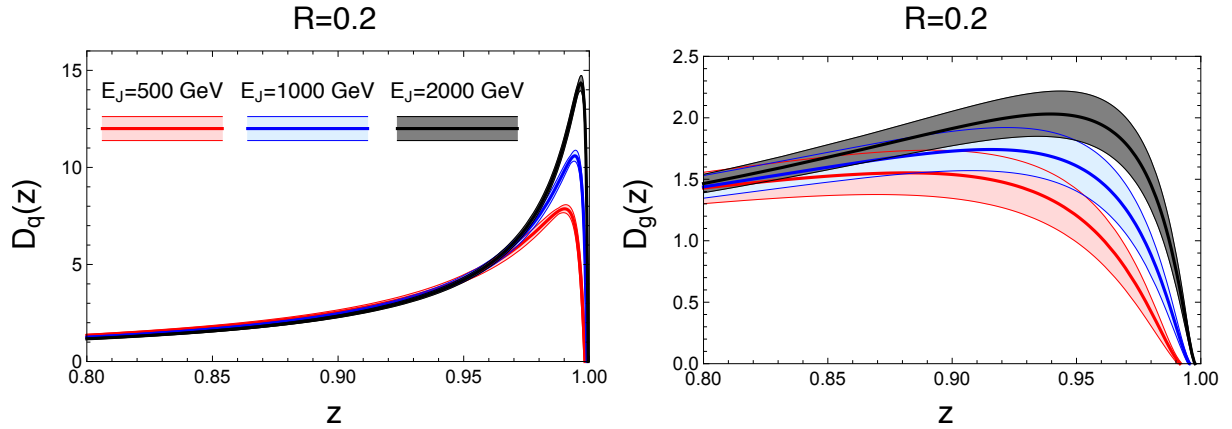


Figure 16:  $D_{J_q/q}(z)$  (left panel) and  $D_{J_g/g}(z)$  (right panel) with different jet energies. Red, blue, and black curves correspond to jet energy  $E_J$  equal to 500, 1000, and 2000 GeV, respectively. Jet radius is chosen as  $R = 0.2$  and the factorization scale is set as  $\mu_f = E_J$ . Error estimation is described in the text.

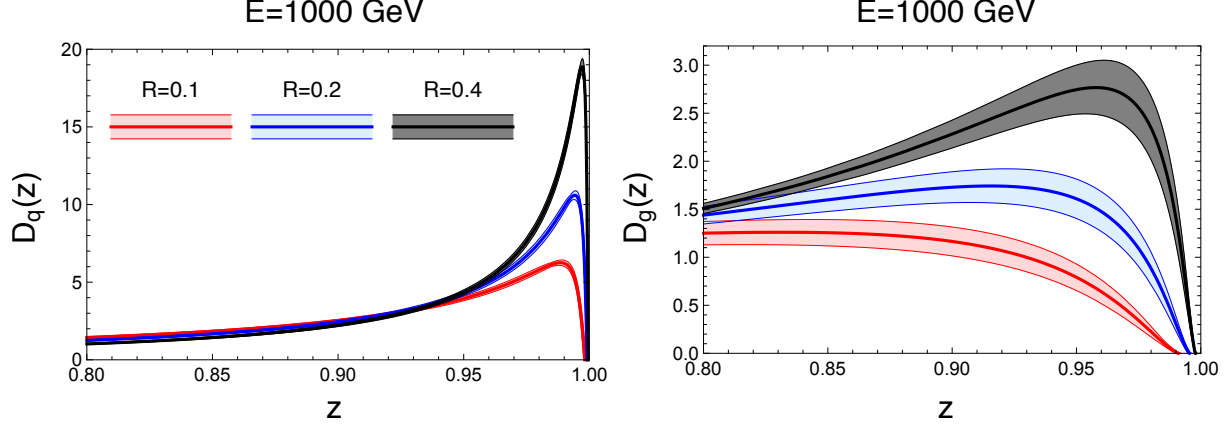


Figure 17:  $D_{J_q/q}(z)$  (left panel) and  $D_{J_g/g}(z)$  (right panel) with different jet radii. Red, blue, and black curves correspond to jet radius  $R$  equal to 0.1, 0.2, and 0.4 respectively. Jet energy are given as  $E_J = 1000$  GeV and the factorization scale is set as  $\mu_f = E_J$ .

In this subsection we show numerical results of the resummed FFJ focusing on the large  $z$  region. As shown in Subsection 3.2.3.1, in order to resum large logarithms in  $D_{J_{k=q,g}/k}(z, \mu_f)$ , the integrated jet functions  $\mathcal{J}_k$  are run from the jet scale  $\mu_c = ER$  to  $\mu_f$ , and the collinear-soft functions  $S_k$  from  $\mu_{cs} = ER(1-z)$  to  $\mu_f$ . Because the FFJ is dependent upon the scale  $\mu_f$  (actually follows DGLAP evolutions.), the shape of the FFJ varies from the choice of  $\mu_f$ . For convenience we choose  $\mu_f = E_J$  throughout this subsection. Error estimations of the jet and the collinear-soft functions are obtained by varying the jet scale and the collinear-soft scale within  $(\mu_c/2, 2\mu_c)$  and  $(\mu_{cs}/2, 2\mu_{cs})$  respectively. Then errors of  $D_{J_k/k}(z, \mu_f)$  are obtained by summing in quadrature.

Based on the factorized expressions in Eqs. (3.92) and (3.95), Fig. 16 shows NLL<sub>G</sub> + NLO results of  $D_{J_q/q}$  and  $D_{J_g/g}$  for different energies of jets with jet radius  $R = 0.2$ . Here NLL<sub>G</sub> represents the NLL accuracy including only global large logarithms from the factorization approach in Subsection 3.2.3.1. For the extreme endpoint region where  $\mu_{cs} = E_J R(1-z) \approx \Lambda_{\text{QCD}}$ , our description is not reliable because of nonperturbative contributions. Fig. 17 shows NLL<sub>G</sub> + NLO results of  $D_{J_q/q}$  and  $D_{J_g/g}$  for different jet radii with the jet energy fixed as 1000 GeV. From Figs. 16 and 17 we can see the tendencies that energetic parton showering

processes are captured more in the jet as the jet energy  $E_J$  and/or the radius  $R$  become larger.

To see the importance of the factorization description on the FFJs, in Fig. 18 we compare the resummed results at  $\text{NLL}_G + \text{NLO}$  and the result using leading DGLAP evolution naively from NLO result in the fixed  $\alpha_s$ . Here using only DGLAP evolution from  $\mu_c = E_J R$  to  $\mu_f = E_J$  can be understood as resumming only large logarithms of  $R$ . As  $z$  goes to 1, the resummed results of only  $\ln R$  become large. However, when we do DGLAP evolution from  $\mu_c = E_J R(1 - z)$  to  $\mu_f = E_J$ , we can see more realistic results. Compared with our factorization approach with accuracy of  $\text{NLL}_G + \text{NLO}$ , both DGLAP evolved results involve much large uncertainties.

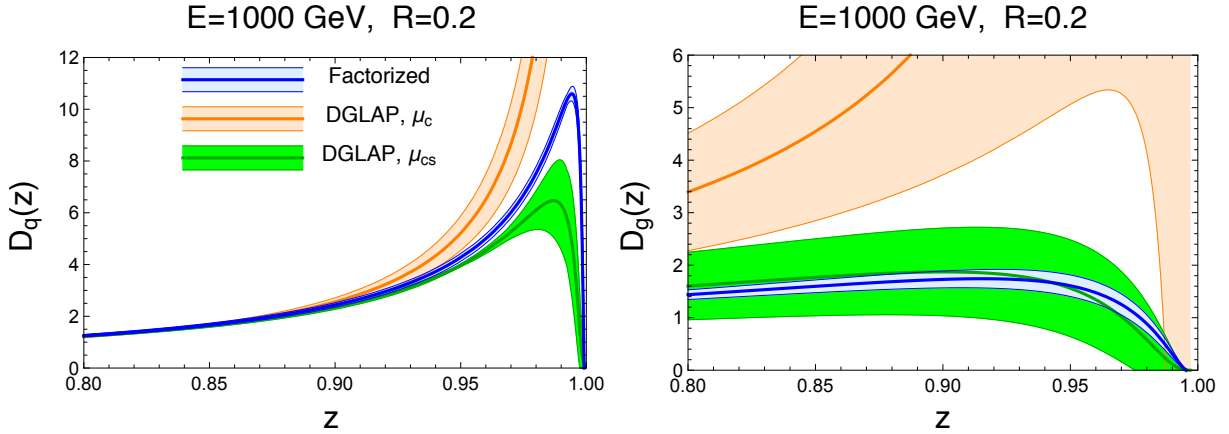


Figure 18: Comparison of the result using leading DGLAP evolution and the resummed result at  $\text{NLL}_G + \text{NLO}$  from the factorization approach. The orange (green) curves are obtained using leading DGLAP evolution with FFJs running from  $\mu_c = E_J R$  ( $\mu_{cs} = E_J R(1 - z)$ ) to  $\mu_f = E_J$ . Blue curves are the resummed result of the FFJs.  $R = 0.2$ ,  $E_J = 1000$  GeV.

Fig. 19 shows the resummed result of the FFJs with accuracy of  $\text{NLL}_{G+\text{NG}} + \text{NLO}$  including leading NGLs discussed in Subsection 3.2.3.2. It can be obtained by multiplication of  $\Delta_{\text{NG}}^{k=q,g}(\mu_c, \mu_{cs})$  in Eq. (3.120) to the FFJ with  $\text{NLL}_G + \text{NLO}$ . The result including leading NGLs gives rise to some suppression to FFJs. The similar suppression can be also seen in the light jet mass distribution for the hemisphere jet production when NLL resummed results with and without NGLs are compared [79]. Because of additional dependences on both  $\mu_c$  and  $\mu_{cs}$  from  $\Delta_{\text{NG}}^{k=q,g}(\mu_c, \mu_{cs})$ , the result with NGLs increases the errors. That might

be reduced if we include the NNLO result at the fixed order in  $\alpha_s$ , which is beyond the scope of our current work.

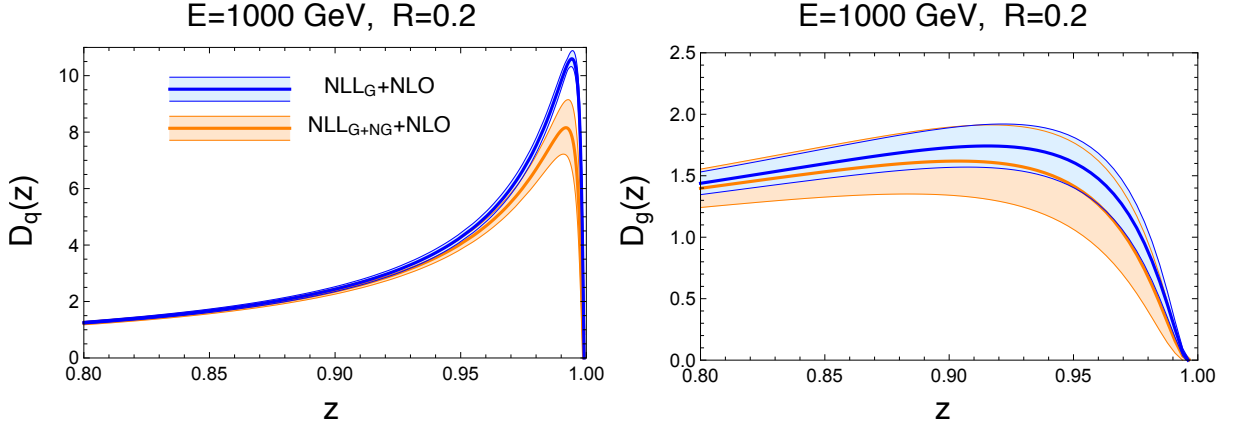


Figure 19: Comparison of the resummed results with (blue) and without resumming the NGLs (orange). Here  $R = 0.2$  and  $E_J = 1000$  GeV.

There is one more comment about error estimations used above. Since  $\mu_{cs} = E_J R(1 - z)$  is  $z$  dependent and bound to hit the Landau pole as  $z \rightarrow 1$ , we have used the following profile function<sup>10</sup> to avoid the Landau pole:

$$\mu_{cs,PF}(z) = \left(1 + \frac{\delta}{1 + \exp[(z - z_1)/(1 - z_1)]}\right) \begin{cases} (1 - z)\mu_c & \text{if } z < z_1 \\ \mu_{\text{Min}} + a(1 - z)^2 & \text{if } z \geq z_1 \end{cases}, \quad (3.122)$$

where  $\mu_{\text{Min}} = 0.3$ ,  $\mu_c = E_J R$ ,  $a$  and  $z_1$  are fixed by requiring that  $\mu_{cs,PF}(z)$  and its first derivative are continuous at  $z = z_1$ .  $\delta = \{0, -0.5, 1\}$  are used for collinear-soft scale variations and the error estimation due to them.  $\mu_{cs,PF}(z)$  is devised to ensure that the collinear-soft scale freezes as it approaches the Landau pole and coincides with  $\mu_{cs}(z)$  otherwise.

### 3.2.5 Conclusion and Outlook

In this section, as shown in Eqs. (3.92) and (3.95), we have developed a factorization theorem of the FFJ with a small  $R$  in the large  $z$  limit. At the scale  $\mu \sim E_J R'$  we first integrate

<sup>10</sup>The profile function provides a smooth transition to non-perturbative regions, and the choice of the form of the function is purely empirical. Ref. to e.g., [83].

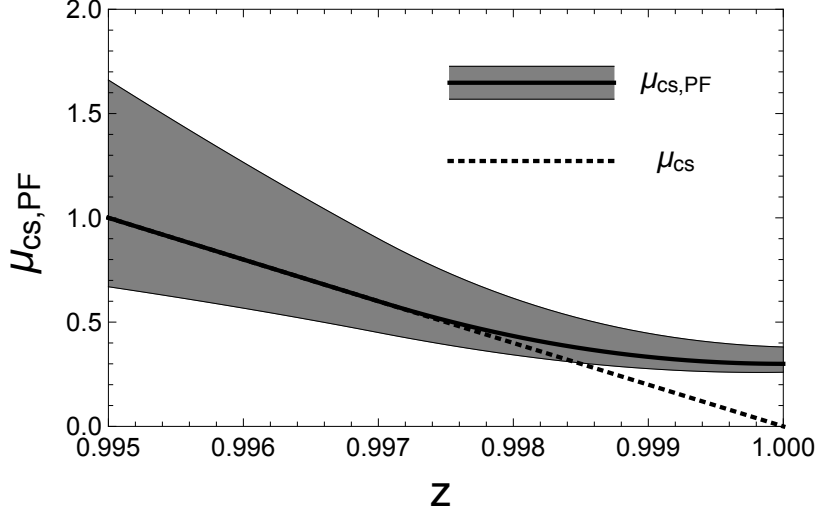


Figure 20: Profile function  $\mu_{cs,PF}$  (solid black curve and gray band, defined in Eq. (3.122) ) is used to estimate errors due to variation of the collinear-soft scale. The dashed line is the  $z$  dependent collinear-soft scale  $\mu_c(1-z)$ .  $\mu_c = 200$  GeV.

out a collinear mode with offshellness  $p_c^2 \sim (E_J R')^2$ , and obtain the integrated jet functions,  $\mathcal{J}_{q,g}$ . At the lower scale  $\mu \sim (1-z)E_J R'$  the collinear-soft mode to read a jet boundary can be active and gives a nonvanishing result at the higher order in  $\alpha_s$ . Combining NLO results of the integrated jet function and the collinear-soft function we can successfully reproduce NLO result of the FFJ in the limit  $z \rightarrow 1$ . The anomalous dimension of each factorized function involves the cusp anomalous dimension, which enables us to systematically resum large logarithms beyond the leading accuracy. As a result we have shown the resummed result to NLL accuracy, which significantly modifies the large  $z$  behavior of the FFJ compared with a naive DGLAP evolution.

We observe that the finite size of the jet radius  $R$  plays an important role in performing successful RG evolution of the FFJ in the large  $z$  limit. Even though  $R$  is small, this makes it possible that an observed jet has a nonzero invariant mass and each factorized function for the FFJ remains to be IR finite. Some similar things happen in the heavy quark fragmentation function (HQFF) in the large  $z$  limit, where the HQFF can be factorized as the heavy quark function and the soft shape function [84, 85]. Due to a nonzero heavy quark mass  $M$ , the



both functions can be IR finite and systematic RG evolution to the scales  $M$  and  $M(1 - z)$  can be done.

Note that the FFJ can be reduced to a light hadron fragmentation function if  $R$  goes to zero. In this case the factorization to collinear and collinear-soft interactions breaks because the relevant anomalous dimensions diverge and RG evolutions become nonperturbative as checked from Eqs. (3.110) and (3.111). A similar result can be applied to the parton distribution function (PDF) near the endpoint. Actually in order to resum large logarithm  $\ln(1 - z)$  in the PDF a similar factorization approach to ours has been considered in Ref. [86], where soft gluon radiations are only responsible for the parton splitting. Interestingly the factorized collinear and soft functions for the PDF contain the rapidity divergences as well as UV divergences. However the rapidity RG evolution turns out to be IR sensitive and become nonperturbative. We also checked if there exist the rapidity divergence in the factorized functions for the FFJ, but the presence of the finite size of  $R$  forbids the rapidity divergences and guarantees ordinary RG evolutions from pure UV divergences.

## 4.0 STUDIES OF QUARKONIUM PRODUCTION IN A JET

This chapter is devoted to the study of quarkonium production in jets. It is based on our work in Refs. [87], [88], and [89]. Each section is self-contained and can be read independently.<sup>1</sup>

In Section 4.1, we use both Monte Carlo simulations and analytic methods to study heavy meson and quarkonium production in jets with certain jet shapes characterized by a parameter called angularity. In Section 4.2, we generalize our work in Section 4.1 to proton collisions and compare our analysis with data measured by the LHCb collaboration. In Section 4.3, we discuss jet energy dependence for quarkonium production in jets and show that measurement of such observables could shed light on quarkonium production mechanisms.

### 4.1 ANALYTIC AND MONTE CARLO STUDIES OF JETS WITH HEAVY MESONS AND QUARKONIA

#### 4.1.1 Introduction

The study of jets and heavy flavor continues to play an important role at the Large Hadron Collider (LHC) and many other high energy and nuclear experiments. Such studies are essential for testing our understanding of Quantum Chromodynamics (QCD) and for calculating backgrounds in searches for new physics. In this section we calculate cross sections for  $e^+e^-$  to jets, where one of the jets contains a hadron with either open or hidden heavy flavor.

---

<sup>1</sup>To avoid confusion, in this chapter (just like the previous chapter), the reference “section” will specifically have two index numbers such as Section 4.1 and Section 4.3 while “subsection” has more than two index numbers such as Subsection 4.1.5 and Subsection 4.3.3.1.

In particular, we will derive factorization theorems and perform analytical Next-to-Leading-Log prime (NLL') resummation<sup>2</sup> for these cross sections using renormalization group (RG) techniques. We will also compare our results with Monte Carlo simulations of the same cross sections.

Recently, there has been considerable interest in cross sections of this type [53, 91, 54, 92, 55, 93, 94, 62, 52, 56]. Ref. [53] demonstrated that the cross section for producing a jet with an identified hadron can be determined using a distribution function called the fragmenting jet function (FJF). FJFs are in turn related to the more commonly studied fragmentation functions (FFs) by a matching calculation at the jet energy scale. This implies that cross sections for jets with an identified hadron provide a new arena to measure FFs, which are more commonly extracted from the semi-inclusive cross section  $e^+e^- \rightarrow H + X$ . Especially important is that this provides an opportunity to extract gluon FFs [52, 56], since quark FFs are more readily studied in  $e^+e^- \rightarrow H + X$ . In addition, it was recently shown in Ref. [62] that since the FFs for quarkonia production can be calculated in the Non-Relativistic Quantum Chromodynamics (NRQCD) factorization formalism [27], FJFs can be used to make novel tests of quarkonium production theory.

The FJF was first introduced in Ref. [53] whose main results can be summarized as follows:

- A factorization theorem for a jet with an identified hadron,  $H$ , is obtained from the factorization theorem for a jet cross section by the replacement

$$J_i(s, \mu) \rightarrow \frac{1}{2(2\pi)^3} \mathcal{G}_i^H(s, z, \mu) dz, \quad (4.1)$$

where  $J_i(s, \mu)$  is the jet function for a jet with invariant mass  $s$  initiated by parton  $i$ , and the renormalization scale is  $\mu$ . The FJF, denoted  $\mathcal{G}_i^H(s, z, \mu)$ , additionally depends on the fraction  $z$  of the jet energy that is carried by the identified hadron. These functions implicitly depend on the jet clustering algorithm and cone size  $R$  used to define the jets. It is also possible to define jet functions and FJFs that depend on the total energy of the jet rather than the invariant mass [55].

---

<sup>2</sup>NLL' includes NLL resummation for each function in the factorization theorem, where all functions are computed to NLO [90].

- The FJFs,  $\mathcal{G}_i^H(s, z, \mu)$ , are related to the well-known FFs,  $D_i^H(z, \mu)$ , by the formulae

$$\mathcal{G}_i^H(s, z, \mu) = \sum_j \int_z^1 \frac{dz'}{z'} \mathcal{J}_{ij}(s, z', \mu) D_j^H(z/z', \mu) + \mathcal{O}(\Lambda_{\text{QCD}}^2/s), \quad (4.2)$$

where the coefficients  $\mathcal{J}_{ij}(s, z, \mu)$  are perturbatively calculable matching coefficients whose large logs are minimized at the jet scale,  $s$ , and are calculated to NLO in Ref. [54]. For heavy quarks the  $\mathcal{J}_{ij}(s, z, \mu)$  have been calculated to  $\mathcal{O}(\alpha_s^2)$  in Ref. [94].

- These matching coefficients obey the sum rule

$$J_i(s, \mu) = \frac{1}{2(2\pi)^3} \sum_j \int_0^1 dz z \mathcal{J}_{ij}(s, z, \mu). \quad (4.3)$$

The properties of FJFs were further studied in Refs. [91, 54, 92, 55, 93]. These papers focused on the FJFs for light hadrons such as pions. FJFs for particles with a single heavy quark were studied in Ref. [94] and FJFs for quarkonia were calculated in Ref. [62].

One important goal of this work is to generalize FJFs to jets in which the angularity is measured. The angularity, denoted  $\tau_a$ , is defined as [95]

$$\tau_a = \frac{1}{\omega} \sum_i (p_i^+)^{1-a/2} (p_i^-)^{a/2}, \quad (4.4)$$

where the sum is over all the particles in the jet, and  $\omega = \sum_i p_i^-$  is the large light-like momentum of the jet. The angularity should be viewed as a generalization of the invariant mass squared of the jet since  $s = \omega^2 \tau_0$ . We calculate the matching coefficients appropriate for jets in which the angularity has been measured, denoted  $\mathcal{J}_{ij}(\tau_a, z, \mu)$ , and verify the  $s \rightarrow \tau_a$  generalization of the sum rules in Eq. (4.3) in Appendix B.2 of this section. The other goal of this work is to study the  $z$  and  $\tau_a$  dependence of the cross section for jets with identified heavy hadrons in  $e^+e^-$  collisions and compare our analytical results to Monte Carlo simulations. We will do this for two-jet events in which  $e^+e^- \rightarrow b\bar{b}$  is followed by fragmentation to  $B$  mesons. We will also study three-jet events with  $e^+e^- \rightarrow b\bar{b}g$  followed by the gluon fragmenting to a jet with a  $J/\psi$ . At the LHC we expect high energy gluons fragmenting to a jet with  $J/\psi$  to be an important production mechanism of  $J/\psi$  at high  $p_T$  and Ref. [62] showed this process is sensitive to the mechanisms underlying quarkonium production. The study in this section will allow comparison of analytic calculations with

Monte Carlo simulations of gluons fragmenting to  $J/\psi$  in jets. In order for this cross section to be physically observable one would either include quarks and antiquarks fragmenting to jets with  $J/\psi$  or one would have to ensure experimentally that the  $J/\psi$  came from the gluon jet in the three-jet event, which could be possible if the other jets are  $b$ -tagged.

In Subsection 4.1.2, we discuss the basics of FJFs for events containing jets where the angularity of the one of the jets is probed. We review various properties of FJFs and their relationship with the more commonly studied FFs. We also present our results for the matching coefficients  $\mathcal{J}_{ij}(\tau_a, z, \mu)$  for jets with measured angularities. Further details of that calculation can be found in Appendix B.2. In Subsection 4.1.3, we present our results for the NLL' cross section for  $e^+e^- \rightarrow 2$  jets where one of the jets contains a  $B$  meson and the angularity of that jet is measured. We find reasonable agreement in both  $z$  and  $\tau_a$  distributions between our analytic calculations and Monte Carlo simulations performed using Madgraph [96] + PYTHIA [97, 98] and Madgraph + HERWIG [99]. In Subsection 4.1.4, we show similar comparisons of analytic versus Monte Carlo calculations for the cross section for  $e^+e^- \rightarrow 3$  jets where one of the jets contains a  $J/\psi$  created via gluon fragmentation. In this case the  $\tau_a$  distributions for the jet are in good agreement, but the Monte Carlo predictions for the  $z$  distributions are inconsistent. We believe that this is due to PYTHIA's modeling of radiation from color-octet states that produces a harder  $z$  distribution than the analytic calculations. In an effort to improve the consistency between NLL' and Monte Carlo calculations, we turn off hadronization in PYTHIA and then convolve the distribution of momenta of the gluons within a jet with the NRQCD color-octet FF at the scale  $2m_c$ . This ad-hoc procedure brings Monte Carlo calculations into much better agreement with analytic NLL' calculations. This suggests that if NRQCD fragmentation could be properly implemented in PYTHIA, consistency with NLL' calculations would be obtained, though more work needs to be done on this problem. In Subsection 4.1.5 we give our conclusions. Appendix B.1 summarizes the renormalization group evolution (RGE) needed for NLL' calculations and also gives the profile functions that are used when computing the scale variation in the NLL' calculations. Appendix B.2 describes the calculation of the matching coefficients and checks that they satisfy the required sum rules that relate them to the jet function.

### 4.1.2 Fragmenting Jet Functions with Angularities

In this subsection we extend the calculation of Ref. [54] to FJFs with measured angularities. We will follow the terminology of Ref. [46], in which a jet whose angularity is measured is referred to as a “measured” jet, while a jet for which only the total energy is measured but the angularity is not is called an “unmeasured” jet. Here we consider the case of two particles as this is the most that will appear in a one-loop calculation. In Ref. [54] the measurement operator in the definition of FJFs forces the mass squared of the jet to be  $s$ . The measurement operator takes the form

$$\delta(\omega(k^+ - l^+ - p^+)) = \delta(s - \omega(l^+ + p^+)), \quad (4.5)$$

where  $k^\mu$  is the parent parton’s momentum and  $l^\mu$  and  $p^\mu$  are the momenta of the partons carrying large lightcone components  $l^- = (1-z)k^-$  and  $p^- = zk^-$  of the parent’s momentum, respectively. The operator definition of the FJF with measured angularities is given by

$$\begin{aligned} \mathcal{G}_i^h(\tau_a, z, \mu) = & \int \frac{dk^+ dp_h^+}{2\pi} \int d^4y e^{-ik^+ y^- / 2} \\ & \times \sum_X \frac{1}{4N_C} \text{tr} \left[ \frac{\not{y}}{2} \langle 0 | \chi_{n,\omega}(y) \delta(\tau_a - \hat{\tau}_a) | Xh \rangle \langle Xh | \bar{\chi}_{n,\omega}(0) | 0 \rangle \right] \end{aligned} \quad (4.6)$$

where at  $\mathcal{O}(\alpha_s)$  the operator  $\hat{\tau}_a$  takes the form (cf. Eq. (4.4))

$$\delta(\tau_a - ((l^+)^{1-a/2} (l^-)^{a/2} - (p^+)^{1-a/2} (p^-)^{a/2}) / \omega). \quad (4.7)$$

Other than replacing Eq. (4.5) with Eq. (4.7), the integrals of all diagrams are the same as in Ref. [54]. However, rather than using the  $\delta$ -regulator and a gluon mass, we will use pure dimensional regularization to regulate all divergences. In this limit, it is possible to show that the one-loop evaluation of the FF yields

$$D_{i \rightarrow j}(z) = \delta_{ij} \delta(1-z) + T_{ij} \frac{\alpha_s}{2\pi} P_{ij}(z) \left( \frac{1}{\epsilon_{UV}} - \frac{1}{\epsilon_{IR}} \right), \quad (4.8)$$

where  $T_{ij}$  are the color structures,  $T_{qq} = C_F$ ,  $T_{gg} = C_A$ ,  $T_{qg} = C_F$ ,  $T_{gq} = T_R$ . Additionally, we have verified that the same  $1/\epsilon_{IR}$  poles appear in the calculation of FJFs and appropriately

cancel in the matching between the FJFs and FFs for all values of  $a < 1$ . This justifies the formula

$$\mathcal{G}_i^h(\tau_a, z, \mu) = \sum_j \int_z^1 \frac{dx}{x} \mathcal{J}_{ij}(\tau_a, x, \mu) D_{j \rightarrow h} \left( \frac{z}{x}, \mu \right), \quad (4.9)$$

which is the analog of Eq. (4.1) for FJFs that depend on the angularities.

Since the matching coefficients  $\mathcal{J}_{ij}(\tau_a, z, \mu)$  are free of IR divergences, we can simplify the matching calculation by using pure dimensional regularization, setting all scaleless integrals to zero and interpreting all  $1/\epsilon$  poles as UV. A detailed calculation of the renormalized finite terms of  $\mathcal{J}_{ij}(\tau_a, z, \mu)$  can be found in Appendix B.2, the results of which are shown below. We parametrize the matching coefficients  $\mathcal{J}_{ij}(\tau_a, z, \mu)$  as

$$\begin{aligned} \frac{\mathcal{J}_{ij}(\tau_a, z, \mu)}{2(2\pi)^3} &= \delta_{ij} \delta(1-z) \delta(\tau_a) \\ &+ T_{ij} \frac{\alpha_s}{2\pi} \left[ c_0^{ij}(z, \mu) \delta(\tau_a) + c_1^{ij}(z, \mu) \left( \frac{1}{\tau_a} \right)_+ + c_2 \delta_{ij} \delta(1-z) \left( \frac{\ln \tau_a}{\tau_a} \right)_+ \right], \end{aligned} \quad (4.10)$$

where

$$\begin{aligned} c_0^{ij}(z, \mu) &= \frac{1-a/2}{1-a} \delta_{ij} \delta(1-z) \left[ \ln^2 \frac{\mu^2}{\omega^2} - \frac{\pi^2}{6} \right] + c^{ij}(z) \\ &- \bar{P}_{ji} \left[ \ln \frac{\mu^2}{\omega^2} + \frac{1}{1-a/2} \ln \left( 1 + \left( \frac{1-z}{z} \right)^{1-a} \right) + (\delta_{ij} - 1) \frac{1-a}{1-a/2} \ln(1-z) \right], \\ c_1^{ij}(z, \mu) &= -\frac{2}{1-a} \delta_{ij} \delta(1-z) \ln \frac{\mu^2}{\omega^2} + \frac{1-a}{1-a/2} \bar{P}_{ij}, \\ c_2 &= \frac{2}{(1-a)(1-a/2)}, \end{aligned} \quad (4.11)$$

with

$$\begin{aligned} c^{qq}(z) &= 1-z + \frac{1-a}{1-a/2} (1+z^2) \left( \frac{\ln(1-z)}{1-z} \right)_+, \\ c^{gg}(z) &= \frac{1-a}{1-a/2} \frac{2(1-z+z^2)^2}{z} \left( \frac{\ln(1-z)}{1-z} \right)_+, \\ c^{qg}(z) &= z, \\ c^{gq}(z) &= 2z(1-z), \end{aligned} \quad (4.12)$$

and where the  $\bar{P}_{ij}$  are the splitting functions of Ref. [54] except for the case  $i = j = q$ ,

$$\begin{aligned}
\bar{P}_{qq} &= P_{qq} - \frac{3}{2}\delta(1-z) = \frac{1+z^2}{(1-z)_+}, \\
\bar{P}_{gg} &= P_{gg} = 2\frac{(1-x+x^2)^2}{x(1-x)_+}, \\
\bar{P}_{qg} &= P_{qg} = x^2 + (1-x)^2, \\
\bar{P}_{gq} &= P_{gq} = \frac{1+(1-x)^2}{x}.
\end{aligned} \tag{4.13}$$

Notice that our results for the matching coefficients  $\mathcal{J}_{ij}(\tau_a, z, \mu)$  are independent of the jet algorithm and the jet size parameter  $R$ . To include modifications of the  $\mathcal{J}_{ij}(\tau_a, z, \mu)$  that come from these effects, one would have to multiply the measurement operator in Eq. (4.7) by an additional  $\Theta$ -function that imposes the phase space constraints required by the jet algorithm. However, for jets with measured angularities, it was shown in Ref. [46] that jet-algorithm dependent terms for cone and  $k_T$ -type algorithms are suppressed by powers of  $\tau_a/R^2$ . Intuitively, this is because as  $\tau_a \rightarrow 0$  all the particles in the jet lie along the jet axis so the result must be insensitive to which algorithm is used and to the value of  $R$  in this limit. For the values of  $\tau_a$  and  $R$  considered in this subsection,  $\tau_a/R^2$  is negligible and we will drop these corrections.

As a non-trivial check of our results we show in Appendix B.2 that our  $\mathcal{J}_{ij}(\tau_a, z, \mu)$  satisfy the following identities and sum rules,

$$\lim_{a \rightarrow 0} \mathcal{J}_{ij}(\tau_a, z, \mu) = \omega^2 \mathcal{J}_{ij}(s, z, \mu), \tag{4.14}$$

and

$$J_i(\tau_a, \mu) = \frac{1}{2(2\pi)^3} \sum_j \int_0^1 dz \, z \, \mathcal{J}_{ij}(\tau_a, z, \mu), \tag{4.15}$$

where  $\mathcal{J}_{ij}(s, z, \mu)$  are the matching coefficients for measured jet invariant mass found in Ref. [54] and  $J_i(\tau_a, \mu)$  are the jet functions for measured jets that can be found in Ref. [46].



### 4.1.3 $e^+e^- \rightarrow 2$ Jets with a $B$ Meson

In this subsection we present an analytic calculation of the cross section for  $e^+e^-$  to two  $b$  jets in which the  $B$  meson is identified in a measured jet. Following the analysis of Ref. [46], the factorization theorem for the cross section for one measured  $b$  jet and one unmeasured  $\bar{b}$  jet is

$$\frac{1}{\sigma_0} \frac{d\sigma}{d\tau_a} = H_2(\mu) \times S^{\text{unmeas}}(\mu) \times J_n^{(\bar{b})}(\mu) \times \left[ S^{\text{meas}}(\tau_a, \mu) \otimes J_n^{(b)}(\tau_a, \mu) \right], \quad (4.16)$$

where  $H_2(\mu)$  is the hard function,  $S^{\text{unmeas}}(\mu)$  and  $S^{\text{meas}}(\tau_a, \mu)$  are the unmeasured and measured soft functions,  $J_n^{(\bar{b})}(\mu)$  is the unmeasured jet function containing the  $\bar{b}$  quark and  $J_n^{(b)}(\tau_a, \mu)$  is the measured jet function containing the  $b$  quark. These describe the short-distance process, surrounding soft radiation, and radiation collinear to unmeasured and measured jets, respectively. At NLO the  $\tau_a$ -independent functions are given by

$$\begin{aligned} H_2(\mu) &= 1 - \frac{\alpha_s(\mu)C_F}{2\pi} \left[ 8 - \frac{7\pi^2}{6} + \ln^2 \frac{\mu^2}{\omega^2} + 3 \ln \frac{\mu^2}{\omega^2} \right], \\ S^{\text{unmeas}}(\mu) &= 1 + \frac{\alpha_s(\mu)C_F}{2\pi} \left[ \ln^2 \frac{\mu^2}{4\Lambda^2} - \ln^2 \frac{\mu^2}{4\Lambda^2 r^2} - \frac{\pi^2}{3} \right], \\ J_n^{(\bar{b})}(\mu) &= 1 + \frac{\alpha_s(\mu)C_F}{2\pi} J_{\text{alg}}^q(\mu), \end{aligned} \quad (4.17)$$

where  $\Lambda$  is a veto on out-of-jet energy,  $r = \tan(R/2)$  and  $J_{\text{alg}}^q(\mu)$  is a function that depends on the algorithm used (and we will use the cone algorithm below) and is given in Eq. (A.18) of Ref. [46]. We note that unlike measured jets, algorithm dependent contributions to the unmeasured jet are not power suppressed. We also note that, beginning at  $\mathcal{O}(\alpha^2)$ , non-global logarithms of the ratio  $Q\tau_a/(2\Lambda r^2)$  begin to appear in the cross-section [38]. For the values of the parameters we consider, these ratios are such that we can treat these logarithms as  $\mathcal{O}(1)$  and thus these would enter as fixed order corrections needed at NNLL' accuracy, which is beyond the scope of this work.

We suppress the dependence of all these functions on scales other than the renormalization scale  $\mu$ . Measured functions are convolved according to

$$f(\tau) \otimes g(\tau) = \int d\tau' f(\tau - \tau') g(\tau'). \quad (4.18)$$

To calculate the differential cross section for a measured jet with an identified  $B$  hadron, we apply the analogous replacement rule in Eq. (4.1) to Eq. (4.16) and use the expression for the FJF in Eq. (4.9) to obtain

$$\frac{1}{\sigma_0} \frac{d\sigma^{(b)}}{d\tau_a dz} = H_2(\mu) \times S^{\text{unmeas}}(\mu) \times J_{\bar{n}}^{(\bar{b})}(\mu) \times \sum_j \left[ \left( S^{\text{meas}}(\tau_a, \mu) \otimes \frac{\mathcal{J}_{bj}^{(b)}(\tau_a, z, \mu)}{2(2\pi)^3} \right) \bullet D_{j \rightarrow B}(z) \right], \quad (4.19)$$

where

$$G(z) \bullet F(z) = F(z) \bullet G(z) \equiv \int_z^1 \frac{dx}{x} F(x) G\left(\frac{z}{x}\right). \quad (4.20)$$

To obtain an NLL' resummed formula for the cross section, we evaluate each function in the factorization theorem in Eq. (4.19) at its ‘‘characteristic’’ scale (where potentially large logarithms are minimized) and, using renormalization group techniques, evolve each function to a common scale,  $\mu$ , which we will choose to be equal to the hard scale. The details of this evolution are discussed in Appendix B.1.

The convolutions in Eq. (4.19) must be performed over angularity over  $S^{\text{meas}}$ ,  $\mathcal{J}_{ij}$ , and factors arising from RG equations. Since such RG factors are distributions ( $\delta$  or plus-distributions) in the angularity our final answer is written in terms of distributions that can be computed analytically using Eqs. (B.18-B.19). Upon performing convolutions and resummation to NLL' accuracy we find for the cross section

$$\begin{aligned} d\sigma(\tau_a, z) &\equiv \frac{1}{\sigma_0} \frac{d\sigma^{(b)}}{d\tau_a dz} = H_2(\mu_H) \times S^{\text{unmeas}}(\mu_\Lambda) \times J_{\bar{n}}^{(\bar{b})}(\mu_{J_{\bar{n}}}) \times \\ &\times \sum_j \left\{ \left( \frac{\Theta(\tau_a)}{\tau_a^{1+\Omega}} \right) \left[ \delta_{bj} \delta(1-z) (1 + f_S(\tau_a, \mu_{S^{\text{meas}}})) + f_{\mathcal{J}}^{bj}(\tau_a, z, \mu_{J_n}) \right] \bullet \frac{D_{j \rightarrow B}(z, \mu_{J_n})}{2(2\pi)^3} \right. \\ &\quad \left. \times \Pi(\mu, \mu_H, \mu_\Lambda, \mu_{J_{\bar{n}}}, \mu_{J_n}, \mu_{S^{\text{meas}}}) \right\}_+, \end{aligned} \quad (4.21)$$

where the ‘+’ distribution is defined in Eq. (B.15) (and acts on all  $\tau_a$ -dependent quantities, including any implicit dependencies arising from the choice of scales  $\mu_F$ ) and  $\Omega(\mu_{J_n}, \mu_{S^{\text{meas}}}) = \omega_{J_n}(\mu, \mu_{J_n}) + \omega_{S^{\text{meas}}}(\mu, \mu_{S^{\text{meas}}})$ , the functions  $\omega_{J_n}$  and  $\omega_{S^{\text{meas}}}$  are given in Appendix B.1, the function  $f_S$  is given by [46]

$$f_S(\tau, \mu) = -\frac{\alpha_s(\mu) C_F}{\pi} \frac{1}{1-a} \left\{ \left[ \ln \frac{\mu \tan^{1-a} \frac{R}{2}}{\omega \tau} + H(-1 - \Omega) \right]^2 + \frac{\pi^2}{6} - \psi^{(1)}(-\Omega) \right\}, \quad (4.22)$$

and  $f_{\mathcal{J}}^{ij}$  are written in terms of the coefficients  $c_0^{ij}$ ,  $c_1^{ij}$  and  $c_2$  presented in Eq. (4.11) as

$$f_{\mathcal{J}}^{ij}(\tau, z, \mu) = T_{ij} \frac{\alpha_s(\mu)}{2\pi} \left( c_0^{ij}(z, \mu) + c_1^{ij}(z, \mu) \left( \ln \tau - H(-1 - \Omega) \right) \right. \\ \left. + c_2 \delta_{ij} \delta(1 - z) \left( \frac{(\ln \tau - H(-1 - \Omega))^2 + \pi^2/6 - \psi^{(1)}(-\Omega)}{2} \right) \right). \quad (4.23)$$

The evolution kernel  $\Pi$  is given in terms of  $K_F(\mu, \mu_0)$  and  $\omega_F(\mu, \mu_0)$  (cf. Appendix B.1),

$$\Pi(\mu, \mu_H, \mu_\Lambda, \mu_{J_n}, \mu_{J_n}, \mu_{S^{\text{meas}}}) = \prod_{F=H, J_n, S^{\text{meas}}} \exp(K_F(\mu, \mu_F)) \left( \frac{\mu_F}{m_F} \right)^{\omega_F(\mu, \mu_F)} \quad (4.24) \\ \times \frac{1}{\Gamma(-\Omega(\mu_{J_n}, \mu_{S^{\text{meas}}}))} \times \prod_{F=J_n, S^{\text{meas}}} \exp(K_F(\mu, \mu_F) + \gamma_E \omega_F(\mu, \mu_F)) \left( \frac{\mu_F}{m_F} \right)^{j_F \omega_F(\mu, \mu_i)},$$

where  $\mu_F$ ,  $m_F$  and  $j_F$  are given in Table 4. Because they involve FFs (cf. Appendix B.2), the  $z$  convolutions must be evaluated numerically. For the fragmentation of the  $b$  quark we use a two-parameter power model FF introduced in Ref. [100], in which  $D_{b \rightarrow B}(z, \mu = m_b = 4.5 \text{ GeV})$  is proportional to  $z^\alpha(1 - z)^\beta$ . Values for the parameters  $\alpha = 16.87$  and  $\beta = 2.628$  with  $\chi_{d.o.f.}^2 = 1.495$  were determined using a fit to LEP data in Ref. [101] for the inclusive process  $e^+e^- \rightarrow B + X$ . Errors in these parameters were not quoted in Ref. [101], so we cannot quantify errors associated with the extracted FF in our calculation. Additionally, we neglect the contribution from the fragmentation of other partons for our  $e^+e^-$  collider studies as in Ref. [101]. In proton-proton collisions at the LHC, gluon FJFs must also be included since the dijet channel  $gg \rightarrow gg$  gives a significant contribution to the production of jets with heavy flavor [56]. For the evolution of the FF up to the jet scale we solve the DGLAP equation using an inverse Mellin transformation as done in Ref. [62].

Fig. 21 shows the  $z$  distributions from  $d\sigma(\tau_0, z)$  for  $\tau_0 = (1.5, 2.0, 2.5) \times 10^{-3}$  of our analytic NLL' calculation (green) and monte-carlo simulations using Madgraph + PYTHIA (black) and Madgraph + HERWIG (red). For each Monte Carlo and for each NLL' calculation, the graphs are independently normalized to unit area. For plots with fixed  $\tau_a$  we use a  $z$ -bin of  $\pm 0.1$  and for plots with fixed  $z$  we use a  $\tau_a$  bin of size  $\pm 2 \times 10^{-4}$ . Jets are reconstructed in PYTHIA using the Seedless-Infrared-Safe Cone (SISCONE) algorithm in the FastJets package [102] with  $R = 0.6$ , which will be used throughout this work. We

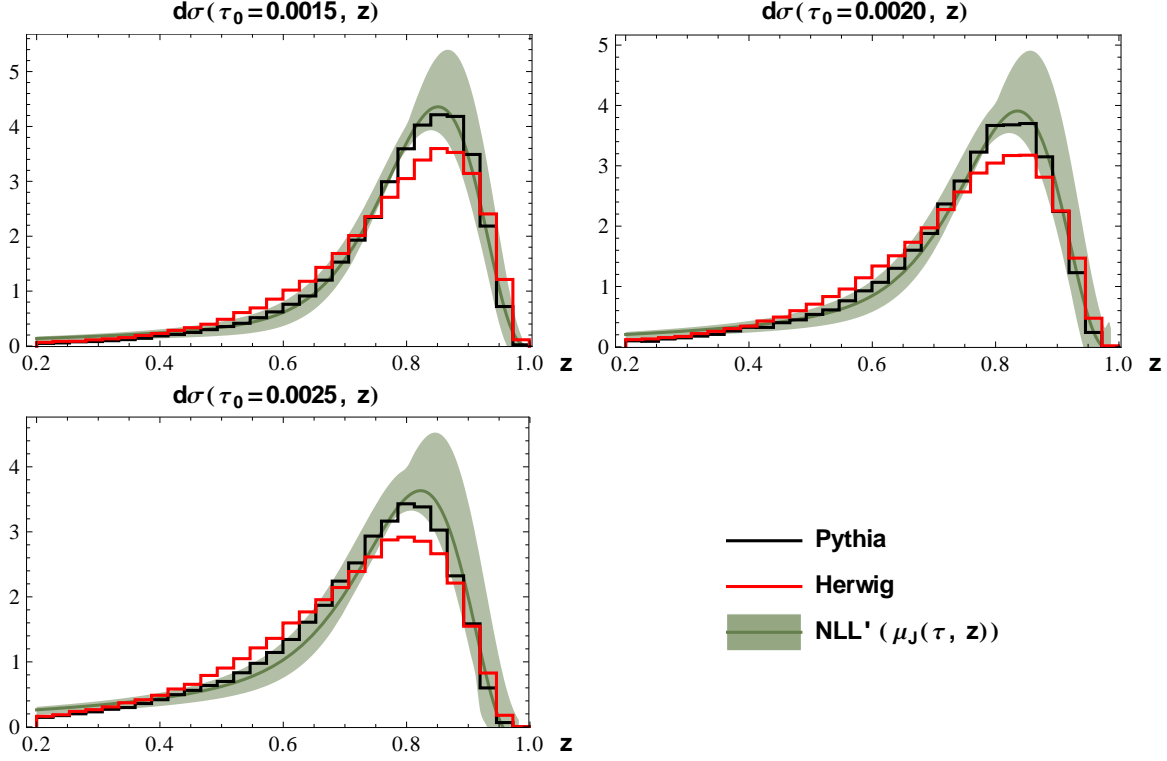


Figure 21: The  $z$  distributions for  $d\sigma(\tau_0, z)$  at  $\tau_0 = (1.5, 2.0, 2.5) \times 10^{-3}$  for analytic calculations with theoretical uncertainty are shown in green. Monte Carlo simulations using Madgraph + PYTHIA and Madgraph + Herwig are shown in black and red, respectively.

produced simulated dijet events at  $E_{cm} = 250$  GeV in which each jet has an energy of at least  $(E_{cm} - \Lambda)/2$  where  $\Lambda = 30$  GeV.<sup>3</sup> The central green line corresponds to the NLL' calculation with the various functions in the factorization theorem evaluated at their characteristic values shown in Table 4, and the green band corresponds to the estimate of theoretical uncertainty obtained by varying the scales of the unmeasured functions by  $\pm 50\%$ , and using profile functions [103, 83, 104] to estimate the uncertainty of the measured functions. Profile functions allow us to introduce an angularity dependent scale variation that freezes at the characteristic scale for high values of  $\tau_a$  where the factorization theorem breaks down and at

<sup>3</sup>This is different than simply placing a cut  $\Lambda$  on energy outside the jets (which is what is assumed in our analytical results), but this difference only appears at  $O(\alpha_s^2)$  in the soft function, which is higher order than we work in this section.

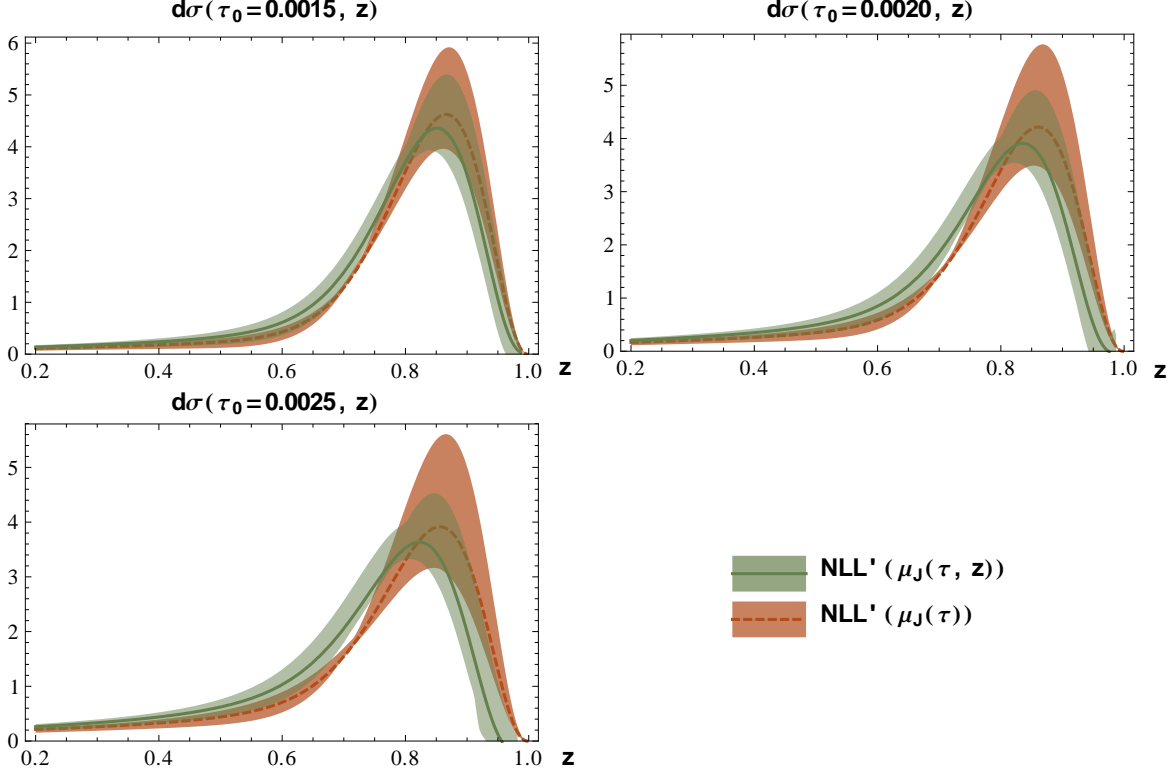


Figure 22: Analytic results for the  $z$  distributions of  $d\sigma(\tau_0, z)$  at  $\tau_0 = (1.5, 2.0, 2.5) \times 10^{-3}$ . The orange curve is calculated with a measured jet scale that does not depend on  $z$  whereas the green curve uses a scale that does depend on  $z$  (as in Fig. 21).

a fixed scale for small values of  $\tau_a$  where we reach the non-perturbative regime. This method for estimating theoretical uncertainties is used throughout this work. Additional details on the profile functions we use can be found in Appendix B.1.

The orange curves in Fig. 22 show the differential cross section as a function of  $z$  for fixed  $\tau_0$  where  $\mu_J(\tau) = \mu_J(\tau, z = 0) = \omega\tau^{1/(2-a)}$  is chosen as the characteristic scale of the measured jet function, and the error band is obtained the same way as for Fig. 21. As in Fig. 21, the green curves show the cross section for a measured jet scale  $\mu_J(\tau, z) = \omega\tau^{1/(2-a)}(1-z)^{(1-a)/(2-a)}$ . The reorganization of logarithms of  $(1-z)$  shown in Eq. (B.22) suggests that we can improve the accuracy of our calculations for  $z \rightarrow 1$  by choosing the characteristic value of the measured jet scale to be  $\mu_J(\tau, z)$ . This improvement is clearly

| Function ( $F$ )  | $H_2$           | $J_{\bar{n}}^b$      | $S^{\text{unmeas}}$ | $\mathcal{J}(\tau, z)$                        | $S^{\text{meas}}(\tau)$   |
|-------------------|-----------------|----------------------|---------------------|---|---------------------------|
| Scale ( $\mu_F$ ) | $E_{\text{cm}}$ | $\omega_{\bar{n}} r$ | $2\Lambda r^{1/2}$  | $\omega_n \tau^{1/(2-a)} (1-z)^{(1-a)/(2-a)}$ | $\omega_n \tau / r^{1-a}$ |
| $m_F$             | $\omega$        | $w_{\bar{n}} r$      | $2\Lambda r^{1/2}$  | $\omega_n$                                    | $\omega_n / r^{1-a}$      |
| $j_F$             | 1               | 1                    | 1                   | $2-a$   | 1                         |

Table 4: Characteristic scales of the different functions in the factorization theorem of Eq. (4.16).

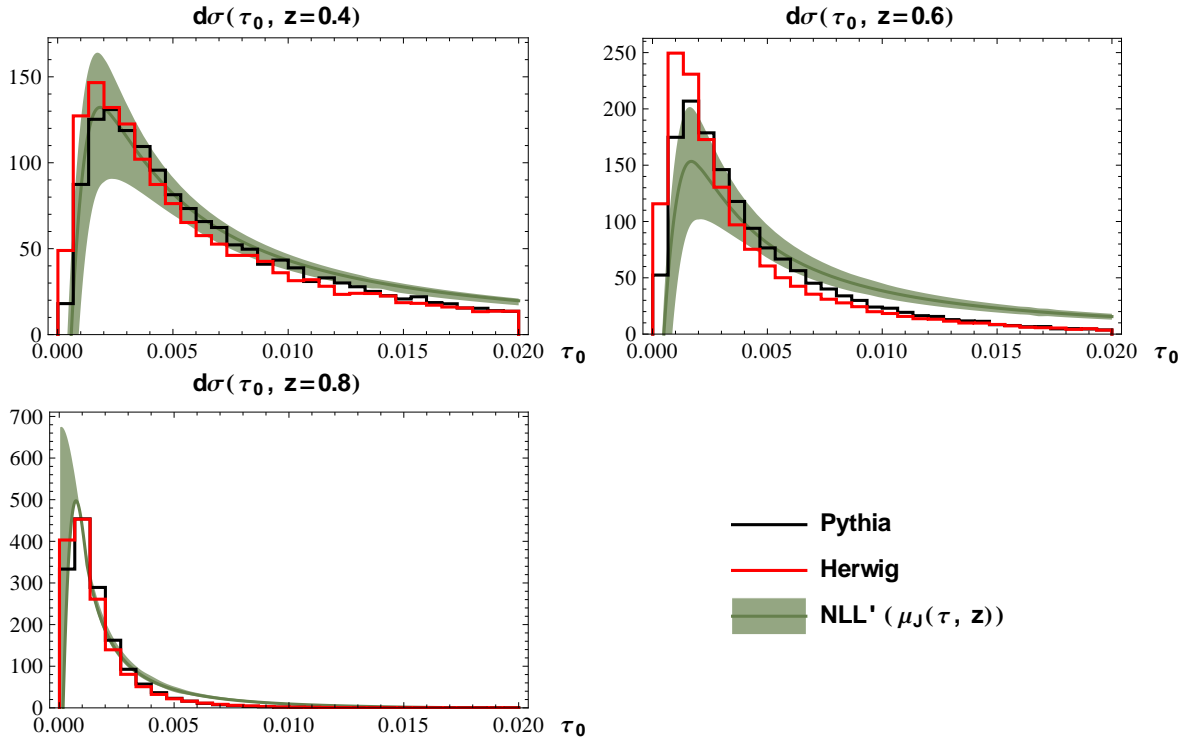


Figure 23: Angularity distributions of  $d\sigma(\tau_a, z)$  for  $a = 0$  at  $z = 0.4, 0.6, 0.8$ . Analytic results are shown as green bands. Monte Carlo results are shown as black lines for Madgraph + PYTHIA and red lines for Madgraph + HERWIG.

seen in Fig. 22 which shows the scale variation for the choices  $\mu_J(\tau)$  and  $\mu_J(\tau, z)$ , the latter choice gives smaller scale variation near the peak in the  $z$  distribution.

In Fig. 23 we present the results for the  $\tau_0$  distributions of the differential cross section  $d\sigma(\tau_a, z)$  for  $z = 0.4, 0.6$ , and  $0.8$ . The color and normalization schemes match those in Fig. 21. We see that for higher values of  $z$  the distributions of  $\tau_0$  are shifted towards smaller values. This is expected since the majority of the energy of the jet is carried by the B meson which results in narrower jets. Figs. 21 and 23 show that our results are consistent within the Monte Carlo uncertainty that is suggested by the difference between PYTHIA and HERWIG predictions. This gives us confidence that the FJF formalism combined with NLL' resummation can be used to correctly calculate both the substructure and the identified hadron's energy fraction within a jet.

#### 4.1.4 $e^+e^- \rightarrow 3$ Jets with the Gluon Jet Fragmenting to $J/\psi$

We can also use the FJF formalism to calculate the cross section for  $e^+e^- \rightarrow 3$  jets with a  $J/\psi$ . As we expect gluon fragmentation to be the dominant production channel at the LHC, we focus on the case where  $J/\psi$  is found within a gluon jet. In addition, we assume that the angularity of this jet is also measured. To obtain a physical observable, one must also include contributions from all jets fragmenting to  $J/\psi$ , however, we expect the contribution from quark jets to be smaller. It is theoretically possible to isolate the  $J/\psi$  coming from gluon jets in experiments by  $b$ -tagging the other two jets in the event, so we will focus on the process  $e^+e^- \rightarrow b\bar{b}g$  followed by gluon fragmentation to  $J/\psi$ .

The analytic expression for this cross section is

$$\begin{aligned} \frac{1}{\sigma_0} \frac{d\sigma^{(g)}}{d\tau_a dz} &= H_3(\mu_H) \times S^{\text{unmeas}}(\mu_\Lambda) \times J_{n_1}^{(\bar{b})}(\mu_{J_{n_1}}) \times J_{n_2}^{(b)}(\mu_{J_{n_2}}) \\ &\times \sum_i \left\{ \left( \frac{\Theta(\tau_a)}{\tau_a^{1+\Omega}} \right) \left[ \delta_{gi} \delta(1-z)(1 + f_S(\tau_a, \mu_{S^{\text{meas}}})) + f_{\mathcal{J}}^{gi}(\tau_a, z, \mu_{J_{n_3}}) \right] \bullet \frac{D_{i \rightarrow J/\psi}(z, \mu_{J_{n_3}})}{2(2\pi)^3} \right. \\ &\quad \left. \times \Pi(\mu, \mu_H, \mu_\Lambda, \mu_{J_{n_1}}, \mu_{J_{n_2}}, \mu_{J_{n_3}}, \mu_{S^{\text{meas}}}) \right\}_+, \end{aligned} \quad (4.25)$$

where  $\Omega \equiv \Omega(\mu_{J_{n_3}}, \mu_{S^{\text{meas}}}) = \omega_{J_n}(\mu, \mu_{J_{n_3}}) + \omega_{S^{\text{meas}}}(\mu, \mu_{S^{\text{meas}}})$ , the  $b$ -quark initiated jets  $J_{n_1}^{(b)}$  and  $J_{n_2}^{(\bar{b})}$  are unmeasured, the expression for  $f_S$  is the same as Eq. (4.22) with  $C_F$  replaced by  $C_A$ , and our expressions for  $f_{\mathcal{J}}^{ij}$  are given in terms of the coefficients  $c_0^{ij}$ ,  $c_1^{ij}$  and  $c_2$  given in Eq. (4.11). Here  $\sigma_0$  is the LO cross section for  $e^+e^- \rightarrow b\bar{b}g$ . We will focus on the Mercedes

Benz configuration in which all three jets have (approximately) the same energy, and consider jets with energies large enough that the mass of  $b$ -quark can be neglected. Here,  $H_3(\mu)$  is  $1 + O(\alpha_s)$  where the  $O(\alpha_s)$  comes from the NLO virtual corrections to  $e^+e^- \rightarrow b\bar{b}g$ . We do not include this correction. The primary effect of its omission will be on the normalization of the cross section, which is not important for our discussion of the distributions we show below, and to increase the scale uncertainty associated with varying  $\mu_H$ ; however this is not a very important source of uncertainty in our calculations.

While the calculation for  $B$  mesons requires a phenomenological FF, the FFs for  $J/\psi$  production can be calculated in NRQCD [27]. Refs. [105, 106, 107, 108] showed that a  $J/\psi$  FF can be calculated in terms of analytically calculable functions of  $\alpha_s(2m_c)$  and  $z$  multiplied by nonperturbative NRQCD long-distance matrix-elements (LDMEs). In  $J/\psi$  production, the most important production mechanisms are the color-singlet mechanism, in which the  $c\bar{c}$  is produced perturbatively in a  $^3S_1^{(1)}$  state, and the color-octet mechanisms, in which the  $c\bar{c}$  is produced perturbatively in a  $^1S_0^{(8)}$ ,  $^3S_1^{(8)}$ , or  $^3P_J^{(8)}$  state. Here  $^{2S+1}L_J^{(1,8)}$  refers to the angular momentum and color quantum numbers of the  $c\bar{c}$ . The numerical values for the corresponding LDMEs are taken to be the central values from the global fits performed in Refs. [2, 3], and are shown in Table 5. The color-singlet LDME scales as  $v^3$ , where  $v$  is the

| $\langle \mathcal{O}^{J/\psi}(^3S_1^{(1)}) \rangle$ | $\langle \mathcal{O}^{J/\psi}(^3S_1^{(8)}) \rangle$ | $\langle \mathcal{O}^{J/\psi}(^1S_0^{(8)}) \rangle$ | $\langle \mathcal{O}^{J/\psi}(^3P_J^{(8)}) \rangle / m_c^2$ |
|---|---|---|---|
| 1.32 GeV <sup>3</sup>                               | $2.24 \times 10^{-3}$ GeV <sup>3</sup>              | $4.97 \times 10^{-2}$ GeV <sup>3</sup>              | $-7.16 \times 10^{-3}$ GeV <sup>3</sup>                     |

Table 5: LDMEs for NRQCD production mechanisms. We use central values taken from global fits in Refs. [2, 3].

typical relative velocity of the  $c\bar{c}$  in the  $J/\psi$ , while the color-octet LDMEs scale as  $v^7$  [27]. This  $v^4$  suppression is clearly seen in the numerical values of the LDMEs in Table 5. In the calculation of the gluon FF, this  $v$  suppression is compensated by powers of  $\alpha_s$  since the leading color-octet contributions are  $O(\alpha_s^2)$  in the  $^1S_0^{(8)}$  and  $^3P_J^{(8)}$  channels and  $O(\alpha_s)$  in the  $^3S_1^{(8)}$  channel, while the color-singlet contribution is  $O(\alpha_s^3)$ . In this work we focus on the gluon FJF,  $\mathcal{G}_g^{J/\psi}$ , and separately compute each of the four NRQCD contributions to  $\mathcal{G}_g^{J/\psi}$ . To calculate  $\mathcal{G}_g^{J/\psi}$ , we evolve each FF from the scale  $\mu = 2m_c$  to the characteristic scale



of the measured jet  $\mu_{J_{n_3}}(\tau_a) = \omega\tau_a^{1/(2-a)}$  using the DGLAP evolution equations. For most values of  $z$  considered in this section, we do not expect that using a  $z$  dependent scale will result in significant improvement in the scale variation. In addition, using a  $z$  dependent scale in the  ${}^3P_J^{(8)}$  channel yields unphysical results, such as negative values for the FF. After evolution, we perform the convolution  $[D \bullet f_J](z)$  in  $z$  with the matching coefficients derived in Subsection 4.1.3.

Before discussing the comparison of our results with Monte Carlo, we briefly review how the Madgraph + PYTHIA Monte Carlo handles color-singlet and color-octet quarkonium production. We produce quarkonia states in Madgraph from the following processes:  $e^+e^- \rightarrow b\bar{b}ggc\bar{c}[{}^3S_1^{(1)}]$ ,  $e^+e^- \rightarrow b\bar{b}gc\bar{c}[{}^1S_0^{(8)}]$ , and  $e^+e^- \rightarrow b\bar{b}c\bar{c}[{}^3S_1^{(8)}]$ . The quantum numbers  ${}^{2S+1}L_J^{(1,8)}$  are for the  $c\bar{c}$  produced in the event. We only include diagrams in which the virtual photon couples to the  $b\bar{b}$  so in all cases the  $c\bar{c}$  plus any additional gluons come from the decay of a virtual gluon. We did not simulate production in the  ${}^3P_J^{(8)}$  channel in  $e^+e^- \rightarrow b\bar{b}g \rightarrow b\bar{b}c\bar{c}g$  because IR divergences in the matrix elements require much longer running times to get the same number of events. We then perform showering and hadronization on these hard processes using PYTHIA. Analysis is done using RIVET [109]. During PYTHIA's showering phase, color-singlet  $J/\psi$  do not radiate gluons. Thus if these  $J/\psi$  are produced within a jet, all surrounding radiation is due to the other colored particles in the event [97, 98]. We require that after showering there are only three jets in the event, two from the  $b$ -quarks and one from a gluon that contains the  $J/\psi$ . We simulate three-jet events at  $E_{cm} = 250$  GeV in the Mercedes-Benz configuration by requiring the jets each have energies  $E_{jet} > (E_{cm} - \Lambda)/3$  with  $\Lambda = 30$  GeV, analagous to what was done in Sec. 4.1.3.

For  $c\bar{c}$  produced in a color-octet state PYTHIA allows the color-octet  $c\bar{c}$  to emit gluons with a splitting function  $2P_{qg}(z)$ . Since  $P_{qg}(z)$  is peaked at  $z = 1$ , the color-octet  $c\bar{c}$  pair typically retains most of its energy after these emissions. This model of the production mechanism is very different than the physical process implied by the NLL' calculation. In the NLL' calculation, the FF is calculated at the scale  $2m_c$ , then evolved up to the jet energy scale using Altarelli-Parisi evolution equations. Since this is a gluon FF, the most important splitting kernel in this evolution is  $P_{gg}(z)$ . We find that the FFs obtained at the jet energy scale are not significantly changed if we use only this evolution kernel and ignore

mixing with quarks. Thus the production process implied by the NLL' calculation is that of a highly energetic gluon produced in the hard process with virtuality of order the jet energy scale, which then showers by emitting gluons until one of the gluons with virtuality of order  $2m_c$  hadronizes into the  $J/\psi$ . Because  $P_{gg}(z)$  is peaked at  $z = 0$  and  $z = 1$  the resulting  $J/\psi$  distribution in  $z$  is much softer than the model employed by PYTHIA. PYTHIA does not allow one to change the actual splitting function, only to modify the color-factor. Therefore, in order to get a softer  $z$  distribution we changed the coefficient of PYTHIA's splitting kernel for a gluon radiating off a color-octet  $c\bar{c}$  pair from  $2P_{qq}$  to  $C_A P_{qq} = 3P_{qq}$ . This results in a slighter softer  $z$  distribution than default PYTHIA, but is still inconsistent with the NLL' calculation. This change does not have significant impact on the  $\tau_a$  distributions. The  $\tau_a$  distributions are generally in better agreement. The variable  $\tau_a$  depends on all of the hadrons in the jet and is therefore less sensitive to the behavior of the  $J/\psi$ , especially when the  $J/\psi$  carries a small fraction of the jet energy. In that case,  $\tau_a$  distributions in the NLL' calculation look similar for all color-octet mechanisms.

In an attempt to see if PYTHIA can be modified to reproduce the  $z$  distributions obtained in our NLL' calculations, and confirm the physical picture of the NLL' calculation described above, we generate  $e^+e^- \rightarrow b\bar{b}g$  events in Madgraph and allow PYTHIA to shower but not hadronize the events. If we allow the shower to evolve to a scale where the typical invariant mass of a gluon is  $2m_c$  and then convolve the gluon distribution with the NRQCD FFs at this scale, we expect that the resulting  $z$  distributions should mimic our NLL' calculation. The lower cutoff scale in PYTHIA's parton shower is set by the parameter `TimeShower:pTmin`, which is related to the minimal virtuality of the particles in the shower, and whose default value is 0.4 GeV. We change this parameter to 1.6 GeV, which corresponds to a virtuality of  $\sim 2m_c$ , then obtain a  $z$  distribution for the gluons by randomly choosing a gluon from the gluon initiated jet. We then numerically convolve this  $z$  distribution with the analytic expression for the NRQCD FF. This procedure, which we will refer to as Gluon Fragmentation Improved PYTHIA (GFIP), yields  $z$  distributions that are consistent with our NLL' result, as we will see below. We tested an analogous procedure for two-jet events with  $B$  mesons by showering  $e^+e^- \rightarrow b\bar{b}$  with PYTHIA with hadronization turned off. We then convolved the resulting  $b$  quark distribution with the  $b$ -quark FF at the scale  $2m_b$ , and found results for  $B$  mesons that

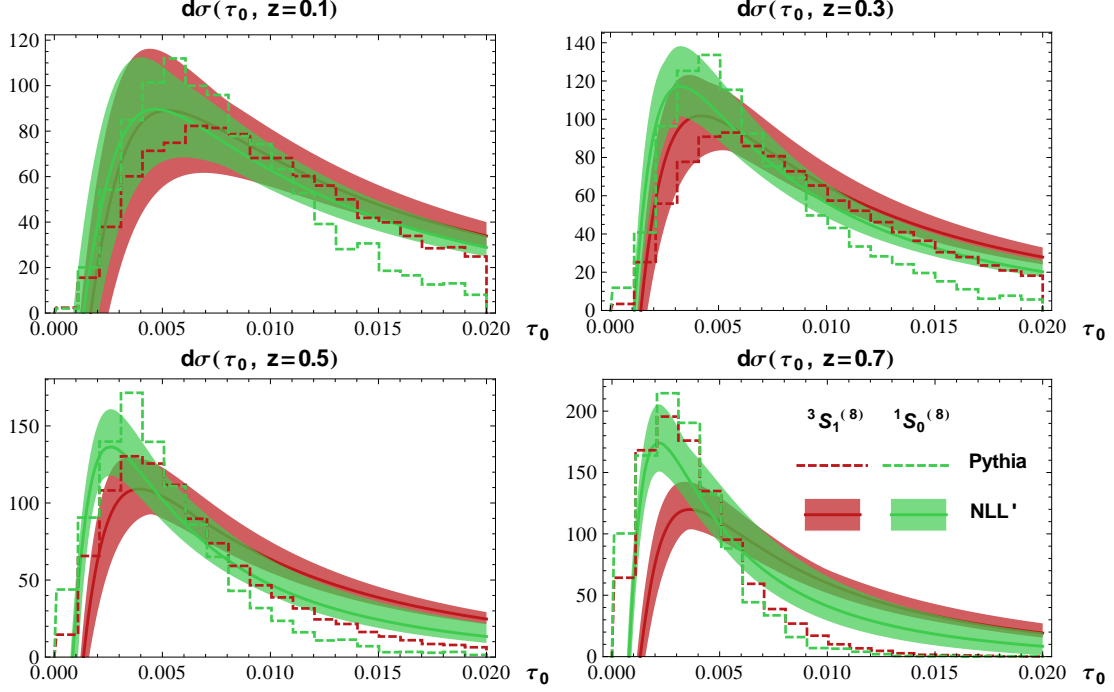


Figure 24: Angularity distributions of  $d\sigma(\tau_a, z)$  for  $a = 0$  at  $z = 0, 1, 0.3, 0.5, 0.7$ . Analytic calculations are shown as red (green) bands for the  $^3S_1^{(8)}$  ( $^1S_0^{(8)}$ ) production mechanisms. Results from Madgraph + PYTHIA are shown as red (green) dashed lines for the same mechanisms.

are consistent with our NLL' calculations. Note that PYTHIA treats the radiation coming from the octet  $c\bar{c}$  pair the same regardless of the angular momentum quantum numbers. In contrast, GFIP like the NLL' calculation gives different results for all three channels by applying different FFs at the end of the parton shower phase. Also GFIP can be applied to all four NRQCD production mechanisms, since convergence issues for the  $^3P_J^{(8)}$  channels are absent.

Fig. 24 shows our NLL' calculation and Madgraph + PYTHIA results for the distribution of  $\tau_0$  for various fixed values of  $z$  for the  $^3S_1^{(8)}$  (red) and  $^1S_0^{(8)}$  (green) channels. We see fairly good agreement between analytic and Monte Carlo results in the peak regions for smaller values of  $z$  and notice some qualitative differences in the tail regions, especially for the  $^1S_0^{(8)}$

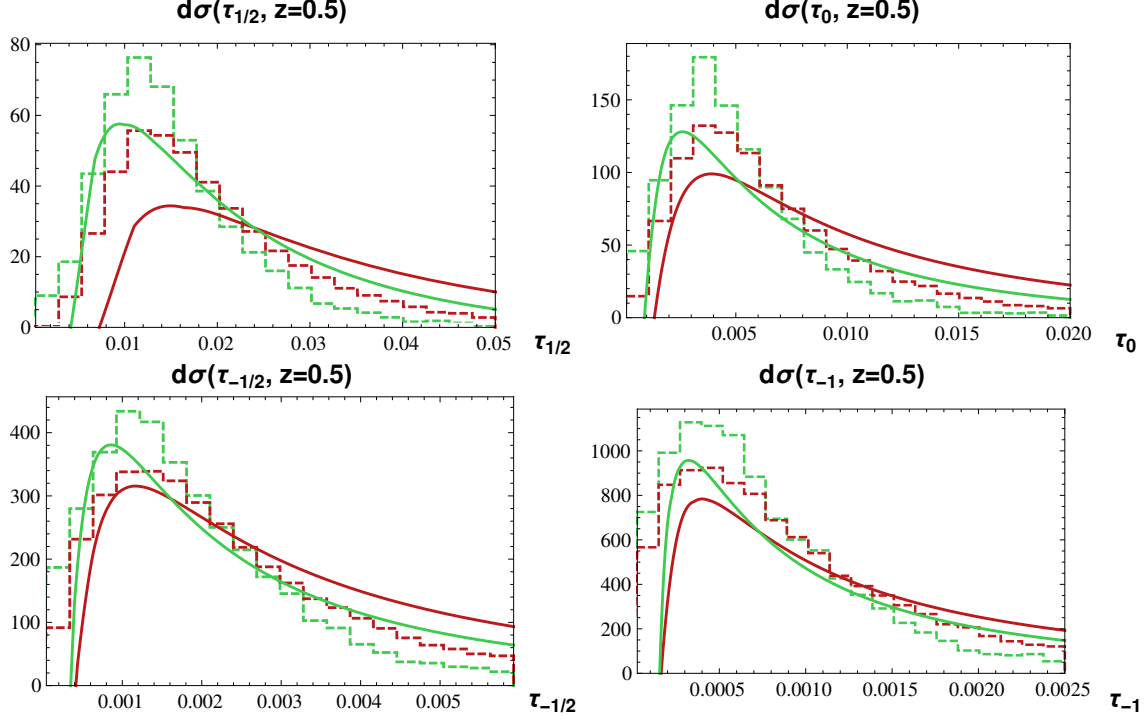


Figure 25: Angularity distributions of  $d\sigma(\tau_a, z)$  for  $a = +1/2, 0, -1/2, -1$  at  $z = 0.5$ . Analytic calculations are shown as red (green) solid lines for the  $^3S_1^{(8)}$  ( $^1S_0^{(8)}$ ) production mechanisms. Results from Madgraph + PYTHIA are shown as red (green) dashed lines for the same mechanisms.

channel. At higher values of  $z$  where the number of final state particles is small, differences in the  $\tau_0$  distributions could be attributed to the increasing influence of Pythia's unrealistic model of quarkonium production. As  $z \rightarrow 0$ , we also see similar  $\tau_0$  dependence for the two color-octet channels in our analytic results. This suggests that in the small  $z$  region, the jet substructure is independent of the production mechanism. Thus, attempts to use angularity distributions to extract the various LDMEs should focus on the range  $0.3 < z < 0.7$ .

In Fig. 25, we show the angularity distributions (without uncertainties) for the  $^1S_0^{(8)}$  and  $^3S_1^{(8)}$  mechanisms for  $a = +1/2, 0, -1/2, -1$ . These are computed analytically and using Monte Carlo and we again see reasonable agreement. As  $a$  is decreased, we see less discrimination between the two production mechanisms. Thus extraction of LDMEs should

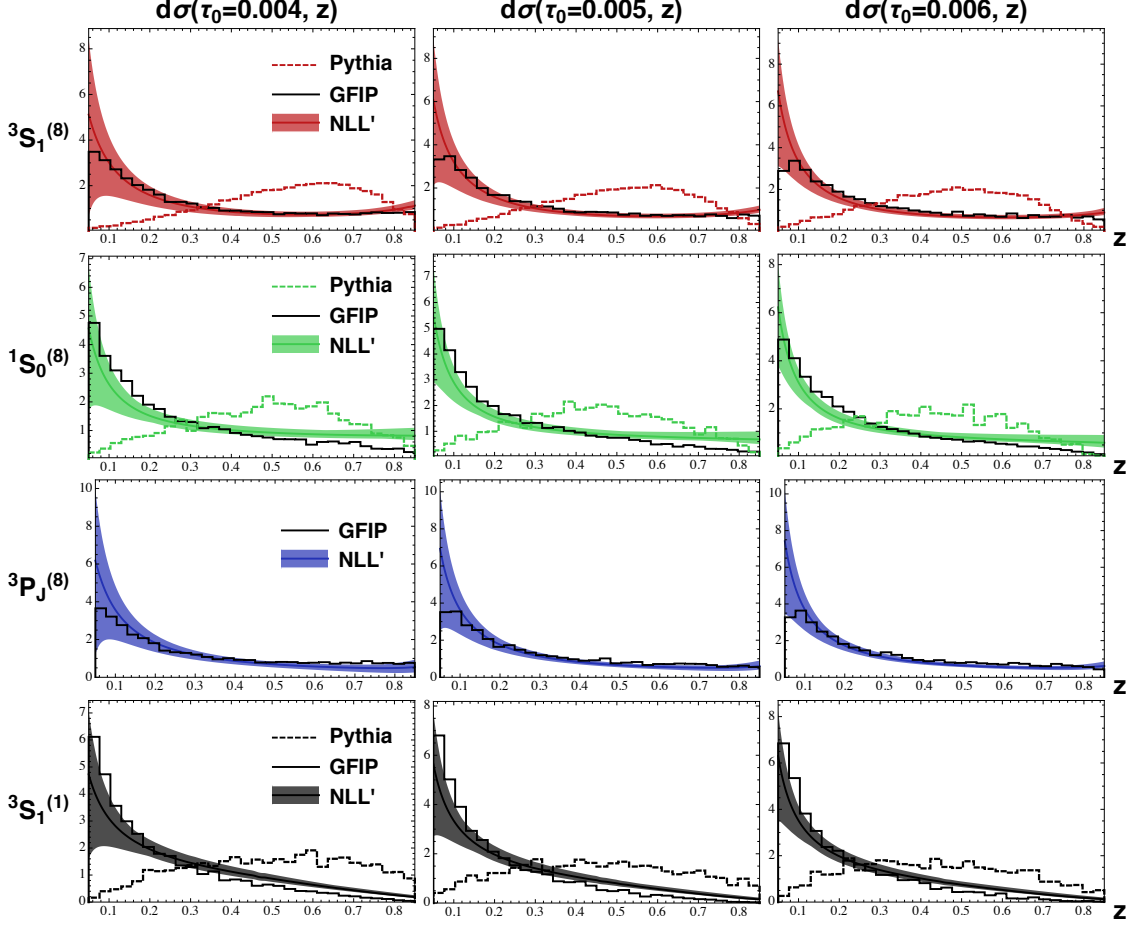


Figure 26:  $z$  distributions of  $d\sigma(\tau_a, z)$  for NLL' analytic calculations (bands), PYTHIA (dashed lines), and GFIP (solid lines) for fixed values of  $\tau_0 = (4, 5, 6) \times 10^{-3}$ .

ideally be done with larger values of  $a$ , for  $a < 1$  where factorization in  $\text{SCET}_I$  holds, with the caveat that there is a trade-off since the predictability of the analytical results is limited for  $a$  too close to 1 since power corrections grow as  $1/(1-a)$  [110].

In contrast to the angularity distributions, Fig. 26 shows that analytic and Monte Carlo calculations of the  $z$  distributions using Madgraph + PYTHIA yield strikingly different results, with Madgraph+PYTHIA yielding a much harder  $z$ -distribution. Fig. 26 also shows the  $z$  distributions using GFIP. The GFIP modification yields significantly different results for the  $z$  distributions that align more closely with NLL' calculation. While this is far from

a proper modification of PYTHIA, it shows us that implementing the missing  $g \rightarrow J/\psi$  fragmentation yields encouraging similarities to our analytical calculations using the FJF formalism with NRQCD FFs. This also suggests that if Monte Carlo is modified to properly include NRQCD FFs at the scale  $2m_c$  it will yield results that are consistent with FJFs combined with NLL' resummation. Correct Monte Carlo implementation of the NRQCD FFs is important because the GFIP modification can only be used to calculate the  $z$  distribution. There are many other jet shape observables, such as  $N$ -subjettiness or  $\Delta R$  (where  $\Delta R$  is the angle between the  $J/\psi$  and the jet axis), that should be able to discriminate between NRQCD production mechanisms, and many of these are most easily predicted using Monte Carlo.

#### 4.1.5 Conclusion

The study of hadrons within jets provides new tests of perturbative QCD dynamics. The distribution in  $z$  (the fraction of jet energy carried by the identified hadron) can be calculated as a convolution of the well-known fragmentation functions (FFs) for that hadron with perturbative matching coefficients that are calculable at the jet energy scale, which is typically well above  $\Lambda_{\text{QCD}}$ . At hadron colliders this provides a new way to extract FFs and will be especially important for pinning down gluon FFs, which are of subleading importance in  $e^+e^-$  colliders where FFs are usually measured. The production of heavy quarkonia within high energy jets in collider experiments also provides new tests of NRQCD.

In this section, we studied cross sections for jets with heavy mesons as a function of  $z$  and the substructure variable angularity,  $\tau_a$ . We provided for the first time the NLO matching coefficients for jets with measured  $\tau_a$ , and used these along with the known RGE for the hard, jet, and soft functions to obtain NLL' accuracy calculations of cross sections for jets with heavy mesons. We considered the production of  $B$  mesons in two-jet events in  $e^+e^-$  collisions at  $E_{\text{cm}} = 250$  GeV as well as  $J/\psi$  production in three-jet events at the same energies. Though not relevant to any experiment, this is useful for comparing NLL' calculations with Monte Carlo simulations of fragmenting jets whose energy is comparable to those measured at the LHC. In the simulations of quarkonia production, the underlying hard

process was generated using Madgraph and then PYTHIA was used to shower and hadronize the events. In the simulations involving B meson production we also used HERWIG.

For  $B$  mesons, we find that the  $z$  and  $\tau_a$  distributions computed using Monte Carlo and NLL' are in excellent agreement, giving us confidence in our analytic approach. In the case of  $J/\psi$ , we considered three-jet events in which the jets all had the same energy and the  $J/\psi$  in both simulation and NLL' calculations was required to come from the gluon jet. This allowed us to study  $J/\psi$  production via the fragmentation of high energy gluon initiated jets, which we expect to be an important mechanism at the LHC. Earlier studies of gluon FJFs in Ref. [62] indicated that the  $z$  and  $E$  dependence of these jets could discriminate between various NRQCD production mechanisms. The analytic NLL' studies of this section are consistent with Ref. [62]; we also find that the  $\tau_a$  and  $z$  distributions can discriminate between different various NRQCD production mechanisms.

For Monte Carlo simulations, we used Madgraph to calculate  $e^+e^- \rightarrow b\bar{b}g$  followed by the gluon fragmenting into a  $c\bar{c}$  pair in either a  $^3S_1^{(8)}$ ,  $^1S_0^{(8)}$ , or  $^3S_1^{(1)}$  state. As explained earlier we do not simulate events in the  $^3P_J^{(8)}$  channel. The events were then showered and hadronized using PYTHIA. While the  $\tau_a$  distributions are similar to analytical calculations, the  $z$  distributions are much harder and their shape looks nothing like the NLL' calculation. We attribute this to a naive model that PYTHIA uses for simulating the radiation of gluons from color-octet  $c\bar{c}$  pairs.

We then considered an alternative simulation approach where  $e^+e^- \rightarrow b\bar{b}g$  events are generated using Madgraph, then PYTHIA is used to shower the event to a low scale near  $2m_c$  without hadronization. The resulting gluon distribution is then convolved with the analytically calculated NRQCD FFs calculated at the scale  $2m_c$ . This procedure yields  $z$  distributions that are in much better agreement with our NLL' calculations.

Future work will focus on extending the NLL' calculations to hadron colliders, where the unmeasured jet and soft function recently calculated in Ref. [104] must be combined with the FJFs of this section. It would be of great interest to compare the results of these calculations with data from the LHC on high energy jets with heavy mesons and quarkonia. Finally, there needs to be more work on improving the understanding of the differences between NLL' and Monte Carlo simulations. Monte Carlo simulations that can properly

simulate the production of quarkonia within jets will be essential for calculating other jet observables for which NLL' calculations are either unavailable or impractical.

## 4.2 NRQCD CONFRONTS LHCb DATA ON QUARKONIUM PRODUCTION WITHIN JETS

The production of quarkonium is a challenging test of Quantum Chromodynamics due to the multiple length scales involved. The LHCb collaboration [111] published the first study of  $J/\psi$  produced within jets. The distribution of the fraction of the jet's transverse momentum,  $p_T$ , carried by the  $J/\psi$ ,  $z(J/\psi)$ , was found to disagree significantly with predictions from the PYTHIA Monte Carlo [97, 98] using leading order calculations of  $J/\psi$  production in the Non-Relativistic Quantum Chromodynamics (NRQCD) factorization formalism [27]. This section provides improved theoretical calculations of the  $z(J/\psi)$  distribution and discusses the implications of the LHCb results for the NRQCD factorization formalism.

Production of quarkonium in hadron colliders has been the subject of experimental and theoretical studies for decades. The problem is challenging because it involves several disparate scales. These include  $p_T$ , which can be much larger than the mass of the bound state,  $\approx 2m_Q$ , where  $m_Q$  is the mass of the heavy quark, as well as scales that are much smaller: the relative momenta,  $m_Q v$  ( $v$  is the typical velocity of the heavy quarks in the bound state), the kinetic energy,  $m_Q v^2$ , and the nonperturbative scale  $\Lambda_{QCD}$ .

The most common approach to calculating quarkonium production is the NRQCD factorization formalism [27]. In this formalism, the cross section for  $J/\psi$  in a  $pp$  collision is written as

$$d\sigma[pp \rightarrow J/\psi X] = \sum_n d\sigma[pp \rightarrow c\bar{c}(n)X] \langle \mathcal{O}^{J/\psi}(n) \rangle,$$

where  $d\sigma[pp \rightarrow c\bar{c}(n)X]$  is the short distance cross section for producing the  $c\bar{c}$  pair in a state  $n$  with definite color and angular momentum quantum numbers and  $\langle \mathcal{O}^{J/\psi}(n) \rangle$  is a long distance matrix element (LDME) that describes the nonperturbative transition of the  $c\bar{c}$  pair in the state  $n$  into a final state containing  $J/\psi$ .  $X$  denotes other possible particles



in the final state. The quantum numbers  $n$  will be denoted  $^{2S+1}L_J^{[i]}$  where the notation for angular momentum is standard and  $i = 1(8)$  for color-singlet (color-octet) states. The short distance cross sections are perturbatively calculable in a power series in  $\alpha_s$ , while the LDMEs are nonperturbative and must be extracted from data. The LDME scale with definite powers of  $v$  so the NRQCD factorization formalism organizes the calculation of quarkonium production (and decay) into a systematic double expansion in  $\alpha_s$  and  $v$ .

For  $J/\psi$  production, the leading matrix element in the  $v$  expansion is  $^3S_1^{[1]}$  which scales as  $v^3$ . The next most important are the color-octet LDMEs:  $^3S_1^{[8]}$ ,  $^1S_0^{[8]}$ , and  $^3P_J^{[8]}$ , which all scale as  $v^7$ .  $J/\psi$  production has been measured in a wide variety of experiments, including  $e^+e^-$ ,  $pp$ ,  $p\bar{p}$ ,  $ep$ ,  $\gamma p$ , and  $\gamma\gamma$  collisions, spanning a wide range of energies. At present, next-to-leading order (NLO) QCD calculations are available for the above mentioned color-singlet and color-octet mechanisms for all these initial states. Global fits to the world's data using these calculations were performed in Refs. [2, 3]. The resulting LDMEs are shown in the first line in Table 6. The LDMEs are consistent with the expected  $v^4$  suppression of the color-octet mechanisms. The global fits in Refs. [2, 3] are reasonably well described by NLO NRQCD, but there are nagging discrepancies that call into question our understanding of quarkonium production. The most notable discrepancy is the polarization puzzle: if the LDMEs of Refs. [2, 3] are used, the produced  $J/\psi$  are predicted to be polarized transverse to their momentum at high  $p_T$ , while in fact they are produced with essentially no polarization. (This is also a problem in  $\Upsilon$  production.) Another important discrepancy is the failure of spin symmetry predictions for  $\eta_c$  production [112, 113]. (For possible solutions to the  $\eta_c$  problem using different extractions of LDMEs see [114, 115, 116].) In light of the failure of NLO QCD to predict the  $J/\psi$  polarization, other authors have proposed alternative approaches to fitting the LDMEs. Refs. [6, 5] have emphasized that NRQCD factorization should be most reliable at the highest values of  $p_T$  and have performed fits that focus exclusively on high  $p_T$   $J/\psi$  production in colliders. Ref. [5] also merges NLO calculations with fragmentation contributions in which Altarelli-Parisi evolution is used to resum logs of  $p_T/m_{J/\psi}$ . The LDMEs from the fits of Refs. [6, 5] are shown in the second and third lines of Table 6, respectively. We will use these three sets of LDMEs in our analysis. There have been other fits to the LDMEs [117, 118] which include explicit feeddown from  $\chi_{cJ}$  states. Since these

|                   | $\langle \mathcal{O}^{J/\psi}(^3S_1^{[1]}) \rangle$<br>$\times \text{GeV}^3$ | $\langle \mathcal{O}^{J/\psi}(^3S_1^{[8]}) \rangle$<br>$\times 10^{-2} \text{GeV}^3$ | $\langle \mathcal{O}^{J/\psi}(^1S_0^{[8]}) \rangle$<br>$\times 10^{-2} \text{GeV}^3$ | $\langle \mathcal{O}^{J/\psi}(^3P_0^{[8]}) \rangle / m_c^2$<br>$\times 10^{-2} \text{GeV}^3$ |
|-------------------|--|--|--|--|
| B & K [2, 3]      | $1.32 \pm 0.20$  | $0.224 \pm 0.59$   | $4.97 \pm 0.44$  | $-0.72 \pm 0.88$   |
| Chao, et al. [6]  | $1.16 \pm 0.20$  | $0.30 \pm 0.12$  | $8.9 \pm 0.98$   | $0.56 \pm 0.21$  |
| Bodwin et al. [5] | $1.32 \pm 0.20$  | $1.1 \pm 1.0$  | $9.9 \pm 2.2$  | $0.49 \pm 0.44$  |

Table 6: LDMEs for NRQCD production mechanisms used in this section in units of  $\text{GeV}^3$ .

effects are not included in our calculations, we do not use these LDME extractions in this work.

Recently, Ref. [62] proposed studying the distribution of quarkonia produced within jets as an alternative test of NRQCD in hadron colliders. Cross sections for jets with identified hadrons are given in terms of fragmenting jet functions (FJF) that were first introduced in Ref. [53] and studied further in Refs. [91, 55, 92, 54, 93, 94, 66, 52, 63, 119, 31, 32]. The FJFs are functions of the jet energy,  $E$ , and the fraction of energy carried by the identified hadron,  $z$ . FJFs are calculable as a convolution of the more inclusive fragmentation function with a perturbative matching coefficient evaluated at the jet energy scale,  $E_J = 2E \tan(R/2)$ . Ref. [62] showed that the quarkonium FJF can be calculated in terms of NRQCD fragmentation functions [107, 106, 105] and that the  $z$  and  $E$  dependence of these cross sections are sensitive to the underlying production mechanisms because NRQCD fragmentations differ for different production mechanisms. For further work see Refs. [87, 64, 56].

In Section 4.1, we used the FJF formalism to compute cross sections for jets with  $B$  mesons and  $J/\psi$  produced within jets in  $e^+e^-$  collisions. For  $B$  mesons we studied  $e^+e^- \rightarrow b\bar{b}$  followed by  $b$  quark fragmenting to a jet with a  $B$  meson. For  $J/\psi$ , we studied  $e^+e^- \rightarrow b\bar{b}g$  followed by gluon fragmentation to a jet with  $J/\psi$ . In both cases we studied the dependence of the cross section on  $z$ , the fraction of the energy carried by the identified hadrons, and the jet's angularity,  $\tau_a$ , [95] a jet substructure variable whose definition can be found in Section

4.1. The analytic expression for these cross sections has the schematic form

$$d\sigma[e^+e^- \rightarrow jets, h] = H \otimes S \otimes J(\otimes J) \otimes \mathcal{G}^h \quad (4.26)$$

where  $h$  is either a  $B$  meson or  $J/\psi$ ,  $\mathcal{G}^h$  is the FJF for the hadron  $h$ ,  $J$  is a jet function for the other jets in the event (there is one  $J$  for  $B$  mesons and two for  $J/\psi$ ),  $S$  is the soft function, and  $H$  is the hard cross section for  $e^+e^- \rightarrow b\bar{b}$  or  $e^+e^- \rightarrow b\bar{b}g$ . Dependence on all kinematic quantities has been suppressed. Evaluation of each of the quantities appearing in Eq. (4.26) shows that they all have logarithms of different scales. The renormalization group equation (RGE) for each of the functions in Eq. (4.26) needs to be solved, and the functions need to be evolved to a common scale so that large logarithms are resummed to all orders in perturbation theory. Details of this formalism for jet cross sections can be found in Ref. [46]. Analytic calculations in Ref. [87] were performed to next-to-leading-log-prime (NLL') accuracy.<sup>4</sup> The calculations were also performed using the PYTHIA Monte Carlo. For  $B$  mesons PYTHIA and NLL' analytical calculations were in good agreement. However, for jets with  $J/\psi$  good agreement was found only in the  $\tau_a$  distributions. The  $z$  distributions predicted by PYTHIA at LO were significantly harder than the  $z$  distributions predicted by the NLL' calculations. This discrepancy between theoretical calculations and PYTHIA is remarkably similar to the discrepancy between data and PYTHIA recently found by the LHCb collaboration. This motivates us to perform calculations similar to those of Ref. [87] to obtain a better description of the LHCb data.

To understand the discrepancy between analytical NLL' calculations of the  $z$  distributions and PYTHIA, one must understand how PYTHIA models the production of quarkonium. In PYTHIA the heavy quark-antiquark pair is produced in the short-distance process in either a color-octet or color-singlet state. If it is in a color-singlet state the heavy quark antiquark pair behaves like a color-singlet particle, emits no gluon radiation, and eventually turns into the quarkonium. If the heavy quark-antiquark pair is produced in a color-octet state, PYTHIA treats the pair as a single colored particle that showers with the splitting function  $2P_{qq}(z)$ . As this splitting function is strongly peaked at  $z = 1$ , the color-octet pair

---

<sup>4</sup>NLL' means that in addition to NLL accuracy,  $O(\alpha_s)$  contributions to the soft, jet, and FJF are also kept.

retains most of its momentum after the shower. At the end of the shower the color-octet quark antiquark emits a soft gluon in order to become a color-singlet quarkonium.

The physical picture of quarkonium production implied by the NLL' analytic calculations in Ref. [87] is quite different from PYTHIA. The FJF that controls the  $z$  dependence of the cross section is, up to  $O(\alpha_s(E_J))$  corrections, equal to the fragmentation function evaluated at the scale  $E_J$ . The evolution of the fragmentation function from the scale  $2m_c$  up to the scale  $E_J$  is governed by Altarelli-Parisi evolution equations. Ref. [87] showed that this is equivalent to producing a hard gluon in the short-distance process with virtuality of order  $E_J$ , allowing that gluon to shower until a gluon with virtuality  $\sim 2m_c$  hadronizes into the  $J/\psi$ . This can be implemented in PYTHIA by simulating events in which the gluon is produced in the hard process, hadronization is turned off, and allowing the parton to shower down to a scale  $\sim 2m_c$ . After this a gluon  $z$  distribution is obtained and convolved manually with a perturbative NRQCD fragmentation function (calculated at lowest order in  $\alpha_s(2m_c)$ ). This procedure was referred to Gluon Fragmentation Improved PYTHIA (GFIP) in Ref. [87], and was shown to give good agreement with the NLL' analytic calculation.

In this section we perform the corresponding calculation for the LHCb experiment using two different methods. Our first method, which we will refer to as GFIP, is analogous to the GFIP calculation of Ref. [87]. We start by generating events corresponding to hard production of  $c$  quarks and gluons in  $pp$  collisions at  $\sqrt{s} = 13$  GeV using MadGraph [96].<sup>5</sup> In the LHCb data, all jets have pseudorapidity  $2.5 < \eta < 4.0$ ,  $R = 0.5$ , and the jets are required to have  $p_T > 20$  GeV. The hard partons generated by MadGraph satisfy the jet constraints of LHCb. PYTHIA is then used to shower the event down to a scale of  $\sim 2m_c$ . Jet algorithms are applied to the output of the PYTHIA shower and the  $c$  quarks and gluons must be within jets of radius  $R = 0.5$  satisfying the criteria of the LHCb data described above. The resulting  $c$  and gluon distributions are shown in Fig. 27. Note that the  $c$  quark distribution is peaked near  $z = 1$  while the gluon  $z$  distribution is much softer and peaked near  $z = 0$ .

The  $p_T$  and  $y$  distributions for the  $c$  quarks and gluons are then convolved manually

---

<sup>5</sup>Contributions to  $J/\psi$  production from quarks other than  $c$  in the hard process are suppressed, either due to soft gluon emission or by  $\alpha_s$  evaluated at a large energy scale. We therefore neglect their contribution.

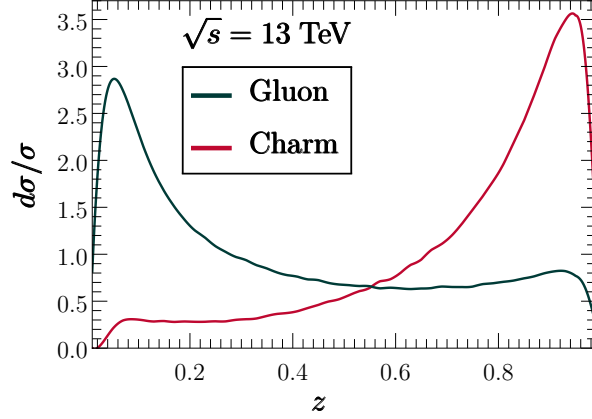


Figure 27: PYTHIA predictions for  $c$  quark and gluon  $z$  distributions (where  $z$  is the fraction of the energy of the parton initiating the jet) after showering to the scale  $2m_c$ .

with the NRQCD fragmentation functions evaluated at leading order (LO) in perturbation theory to obtain  $p_T$  and  $y$  distributions for  $J/\psi$ . For gluons we include  $^3S_1^{[1]}$ ,  $^3S_1^{[8]}$ ,  $^1S_0^{[8]}$  and  $^3P_J^{[8]}$  fragmentation functions, because the  $v^4$  suppression of the color-octet LDMEs is compensated by powers of  $\alpha_s$  for  $^1S_0^{[8]}$  and  $^3P_J^{[8]}$  and  $\alpha_s^2$  for  $^3S_1^{[8]}$ . See Ref. [62] for the explicit expressions for the LO NRQCD fragmentation functions. Color-singlet and color-octet fragmentation functions start at the same order in  $\alpha_s$  for charm quarks so we include only color-singlet fragmentation for charm quarks. LHCb requires both muons have  $2.0 < \eta < 4.5$ ,  $p > 5$  GeV, and  $p_T > 0.5$  GeV. The energy cut clearly suppresses contributions from partons with low  $z$  and hence enhances the contribution from  $c$  quark initiated jets. We implement the muon cuts by assuming the  $J/\psi$  are unpolarized and therefore decays to  $\mu^+\mu^-$  isotropically in its rest frame, and the LHCb cuts on the muons are applied to the muons after they are boosted back to the lab frame. From this a normalized distribution in  $z(J/\psi)$  is constructed for each production mechanism. Each mechanism is characterized by an initial parton  $i$  and quantum numbers  $n$ , and is multiplied by a weight

$$r(i, n) = \frac{d\hat{\sigma}(pp \rightarrow i + X) \int_0^1 dz D_{i \rightarrow J/\psi}^n(z)}{d\hat{\sigma}(pp \rightarrow c + X) \int_0^1 dz D_{c \rightarrow J/\psi}^{^3S_1^{[1]}}(z)}. \quad (4.27)$$

The weight in Eq. (4.27) ensures that the total number of  $J/\psi$  coming from each mechanism are in the proper ratio where  $D_{i \rightarrow J/\psi}^n(z)$  are calculated at the scale  $2m_c$ . This is where the

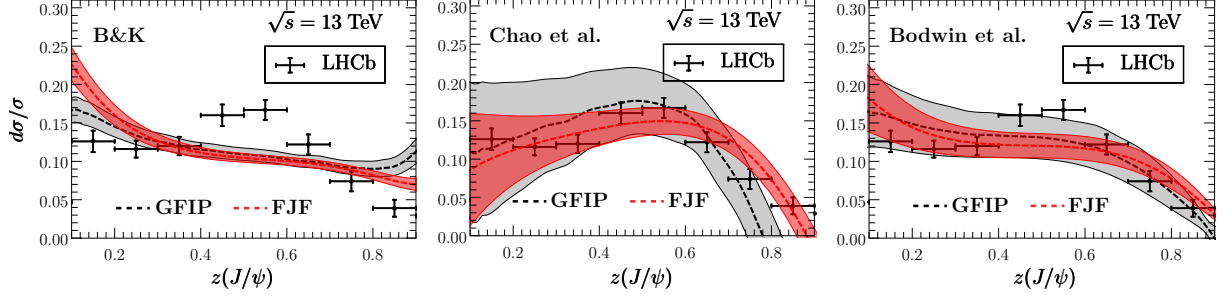


Figure 28: Predicted  $z(J/\psi)$  distribution using GFIP (gray) and FJF (red) for the three choices of LDME in Table 1 and the LHCb measurements of  $z(J/\psi)$ .

fitted LDMEs enter the calculation as  $D_{i \rightarrow J/\psi}^n(z) \propto \langle \mathcal{O}^{J/\psi}(n) \rangle$ . The LHCb data is normalized so that the sum of the heights of the bins adds to 1. Because of possible large corrections near  $z \rightarrow 0$  and  $z \rightarrow 1$ , we only compare with LHCb data in the range  $0.1 < z < 0.9$  and normalize our distributions to the sum of the data in these bins.

Our second method, which we refer to as the FJF method, employs FJFs combined with hard events generated by Madgraph at LO. In calculating the FJFs, logarithms of  $m_{J/\psi}/E_J$  are resummed using leading order DGLAP equations to evolve the fragmentation functions from the scale  $2m_c$  to the jet energy scale,  $E_J$ . Madgraph calculates the remaining terms in the factorization theorem to LO in perturbation theory. This does not include NLL' resummation for the remaining terms in the factorization theorem, however the  $z(J/\psi)$  dependence of the cross section is controlled primarily by the FJF. The energy distribution of hard partons is combined with the FJFs for anti- $k_T$  jets [120] with  $R = 0.5$  to produce a  $z(J/\psi)$  distribution for each of the five mechanisms. From the GFIP calculations, we know as a function of  $z$  the fraction of  $J/\psi$  that survive the muon cut and we apply this correction to our analytic calculations. The  $z(J/\psi)$  distributions from each mechanism are weighted by the factors in Eq. (4.27) as before. The FJF is appropriate for  $n$ -jet cross sections like Eq. (1). Inclusive FJFs [63, 119, 31, 32] differ by a contribution from out-of-jet radiation that is power suppressed for  $R \sim O(1)$  [46].

Fig. 28 shows the predicted  $z(J/\psi)$  distributions for the three choices of LDME's in Table 6 using the GFIP (gray) and FJF (red) methods, which are in good agreement. Uncertainties

are due to the LDMEs only. In the case of Ref. [5], the errors in Table 6 are supplemented with an error correlation matrix [121]. In Ref. [6] a fixed relationship between the  $^3S_1^{[8]}$  and  $^3P_J^{[8]}$  LDMEs is required to obtain unpolarized  $J/\psi$ . This constraint is taken into account when computing the uncertainty due to the LDMEs. These constraints significantly reduce the uncertainty in the predictions relative to naively adding uncertainties in Table 6 in quadrature. Other sources of uncertainty such as scale variation have not been included. Estimating theory uncertainties reliably in the absence of a complete factorization theorem is difficult. For example, using the FJF method, the  $\mu$  dependence of the FJF should be cancelled by  $\mu$  dependence in hard and soft functions that have not been computed. Note that since the normalization of theoretical curves is fixed to the LHCb data, any scale variation that affects normalization but not the shapes of the  $z(J/\psi)$  distribution will not contribute to the uncertainty. Especially at low values of  $z$ , the underlying event and double parton scattering give additional theoretical uncertainties. However, it is not clear how estimate these uncertainties.

All three choices of LDMEs give better agreement to the LHCb data than default PYTHIA shown in Ref. [111]. This gives support to the picture of quarkonium production in Section 4.1 and this section. The LDMEs from global fits [2, 3] give worse agreement than the fits from Refs. [5, 6]. The LHCb data is a decreasing function of  $z(J/\psi)$  as  $z(J/\psi) \rightarrow 1$ . This is a property of the  $^3S_1^{[1]}$  and  $^1S_0^{[8]}$  FJFs, but not the  $^3S_1^{[8]}$  and  $^3P_J^{[8]}$  FJFs, which actually diverge as  $z \rightarrow 1$ . In order to obtain negligible polarization at high  $p_T$ , the  $^3S_1^{[8]}$  and  $^3P_J^{[8]}$  LDMEs of Refs. [5, 6] have relative signs such that they roughly cancel, so the  $^1S_0^{[8]}$  dominates production and  $J/\psi$  are unpolarized. The same cancellation here allows the  $z(J/\psi)$  distribution go to zero as  $z(J/\psi) \rightarrow 1$ . Such a cancellation does not occur for the LDMEs from the global fits so the  $z(J/\psi)$  distribution starts to turn up at large  $z(J/\psi)$ .

To summarize, we have analyzed the recent LHCb data on  $J/\psi$  production within jets. We used a combination of Madgraph, PYTHIA, and LO NRQCD fragmentation functions first introduced in Ref. [87] as well as an approach based on Monte Carlo evaluation of the hard process combined with  $J/\psi$  FJFs evaluated at the jet energy scale. Both methods yield  $z(J/\psi)$  distributions that agree much better with data than default PYTHIA simulations. The  $z(J/\psi)$  distributions are very well described by LDMEs from fits to large  $p_T$  data, and

less well described by LDMEs from global fits. It would be interesting to perform a combined fit to the LHCb data and the large  $p_T$  data used in Refs. [5, 6] to see if consistent LDMEs with smaller errors can be obtained. Experimental measurement of jets at central rapidity and the polarization of  $J/\psi$  as a function of  $z(J/\psi)$  [122] would also be of interest. Finally it would be especially interesting to find ways of discriminating charm and gluon initiated jets [123], as a sample containing only gluon initiated jets will have greater sensitivity to color-octet LDMEs.

### 4.3 QUARKONIUM POLARIZATION AND THE LONG DISTANCE MATRIX ELEMENTS HIERARCHIES USING JET SUBSTRUCTURE

#### 4.3.1 Introduction

Analyzing quarkonium production in jets provides a new way of probing the physics involved in their production. Recent developments include the LHCb measurements of  $J/\psi$  production in jets [111] and the related analyses [124, 88, 125]. A factorization theorem based on Non-Relativistic QCD (NRQCD)<sup>6</sup> can be used to calculate the cross section for  $J/\psi$  production [26, 126]. Due to the large mass of the charm quark ( $m_c$ ), the short distance production of the  $c\bar{c}$  pair can be calculated perturbatively while the non-perturbative physics of the hadronization into a  $J/\psi$  is captured by the long distance matrix elements (LDMEs) of the relevant production channels ( $^1S_0^{[8]}$ ,  $^3S_1^{[8]}$ ,  $^3P_J^{[8]}$ , and  $^3S_1^{[1]}$ ). The predictive power of the theory is then predicated on our knowledge of these LDMEs. Different groups have extracted these matrix elements by using various fits to the data [5, 2, 3, 6] but have arrived at very different values. Currently the NRQCD factorization theorem can consistently fit the unpolarized  $J/\psi$  production cross section [129].

The  $c\bar{c}$  pair produced by the fragmentation of a nearly on-shell gluon<sup>7</sup> should inherit the

---

<sup>6</sup>NRQCD is an effective theory with a double expansion in the relative velocity  $v$  of the heavy quark and anti-quark bound state and the strong coupling constant  $\alpha_s$  [126, 26, 127, 128].

<sup>7</sup>For  $J/\psi$  production via gluon fragmentation in NRQCD, the  $^3S_1^{[1]}$  contribution is leading order in the  $v$  expansion since the color octet channels are suppressed by  $v^4$ . But the  $^3S_1^{[1]}$  is suppressed relative to the  $^3S_1^{[8]}$  channel by power of  $\alpha_s^2$ . The matching onto  $^3P_J^{[8]}$  and  $^1S_0^{[8]}$  is down by  $\alpha_s$  compared to  $^3S_1^{[8]}$  but their



transverse polarization of the gluon. Due to the spin symmetry of the leading order NRQCD Lagrangian, this polarization remains intact during the non-perturbative hadronization process (up to power corrections) [131, 132]. At leading order in  $\alpha_s$ , only the  $^3S_1^{[8]}$  channel for the gluon contributes among the octet channels and since the color octet contribution is expected to dominate at high  $p_T$  [133], the  $J/\psi$  meson should be produced with significant polarization at high  $p_T$ . However this prediction of NRQCD is at odds with the measurements of the  $J/\psi$  polarization [134, 135, 136]. Understanding this polarization puzzle is one of the most important challenges in quarkonium physics [15].

A method based on jet substructure techniques to study the different production mechanisms of the  $J/\psi$  was proposed in Ref. [62]. By using the properties of the Fragmenting Jet Functions (FJF) [53], it is predicted in Ref. [62] that for a jet of energy  $E$  and cone size  $R$ , containing a  $J/\psi$  with energy fraction  $z$  ( $z = E_{J/\psi}/E$ ), if the FJF is a decreasing function of the jet energy, then the dominant contribution to the  $J/\psi$  production at high  $p_T$  should be the depolarizing  $^1S_0^{[8]}$  channel and hence, if confirmed by the data, this would resolve the polarization puzzle.

In this work, we investigate how the predictions of the diagnostic tool introduced in Ref. [62] are affected by inclusion of the hard scattering effects. To do this, we calculate the total production cross section for the  $J/\psi$ . This should make the comparison of theory with experiments much simpler since the cross section can be directly measured. In order to make the distinction between various production channels, we calculate the cross section normalized in two different ways. In one case we normalize by summing over the contribution of all the channels and integrating over  $z$  while in the other case we normalize by using the 1-jet inclusive cross section. Additionally we also make comparisons between the LDMEs extracted by various groups.

The main result of this section is that the prediction made in Ref. [62], regarding the shapes of the FJF's, is also true for the cross section. By using a combination of differently normalized cross sections, we can break the degeneracy of the production channels and isolate the dominant contribution to the  $J/\psi$  production at high  $p_T$ . Our results show that if the

---

LDMEs are of the same order as  $^3S_1^{[8]}$  in  $v$ . An alternate power counting for charmonium production is formulated in Ref. [130].

normalized cross section is a decreasing function of the jet energy at large  $z$ , in particular for  $z > 0.5$ , then the  $^1S_0^{[8]}$  channel dominates at high  $p_T$  and this prediction should be easily verifiable with the LHC data. A recent work [124] also proposed using observables similar to ours to probe the  $J/\psi$  production mechanisms.<sup>8</sup>

### 4.3.2 The Fragmenting Jet Functions

We briefly review the factorization theorem for the production of  $J/\psi$  [137, 91, 54, 92, 55, 93] before moving onto our main results in the next subsection. We consider the process  $pp \rightarrow$  dijets at  $\sqrt{s} = 13$  TeV and integrate over one of the jets, assuming that the other jet contains an identified  $J/\psi$ . The dijet cross section [137] with one jet of energy  $E$ , cone size  $R$  and a  $J/\psi$  in the jet carrying an energy fraction  $z$ , is schematically of the form

$$\frac{d\sigma}{dEdz} = \sum_{a,b,i,j} H_{ab \rightarrow ij} \otimes f_{a/p} \otimes f_{b/p} \otimes J_j \otimes S \otimes \mathcal{G}_i^\psi(E, R, z, \mu), \quad (4.28)$$

where  $H_{ab \rightarrow ij}$  is the hard process,  $f_{a/p}$  and  $f_{b/p}$  are the parton distribution functions (PDF),  $J_j$  is the jet function for the jet not containing the  $J/\psi$ ,  $S$  is the soft function and  $\mathcal{G}_i^\psi(E, R, z, \mu)$  is the FJF for the jet containing the  $J/\psi$ . The parton  $i$  can be a gluon, charm or an anti-charm (contributions of the other partons are suppressed). We are interested in the  $E$  and  $z$  dependence of the cross section, which comes from the hard function (including PDFs) and the FJF. We integrate over the jet originating from the parton  $j$  so the jet function  $J_j$  enters the cross section multiplicatively. The soft function  $S$  does not affect  $\mathcal{G}_i^\psi(E, R, z, \mu)$ ,  $R$ ,  $E$  and  $z$  (up to power corrections) [62] and so it also enters the cross section multiplicatively. Hence both the jet function  $J_j$  and the soft function  $S$  give an overall normalization to the cross section and are ignored in the rest of our analysis. In Ref. [62], the hard function was not included but here we calculate the normalized cross section, including both the charm quark and gluon contributions, and account for its  $E$  dependence.

---

<sup>8</sup>Ref. [124] differentiates between the NRQCD global fits based on inclusive  $J/\psi$  cross section and suggests using the polarization measurements of  $J/\psi$  meson produced in the jets as a way of constraining the heavy quarkonium production mechanisms.

The FJF can be further factorized [137] into perturbatively calculable coefficients  $\mathcal{J}_{ij}(E, R, z, \mu)$  and the fragmentation function  $D_{j \rightarrow \psi}$ :

$$\mathcal{G}_i^\psi(E, R, z, \mu) = \int_z^1 \frac{dy}{y} \mathcal{J}_{ij}(E, R, y, \mu) D_{j \rightarrow \psi}\left(\frac{z}{y}, \mu\right) \left(1 + \mathcal{O}\left(\frac{m_\psi^2}{4E^2 \tan^2(R/2)}\right)\right). \quad (4.29)$$

The collection of NRQCD based fragmentation functions  $D_{j \rightarrow \psi}$  used in this section can be found in Ref. [62].

Large logarithms in  $\mathcal{J}_{ij}(E, R, z, \mu)$  are minimized at the scale  $\mu = 2E \tan(R/2)(1 - z)$  and can be easily resummed using the jet anomalous dimension [55]. But we do not consider this resummation in this work since for us,  $1 - z \sim \mathcal{O}(1)$  [62]. Instead we evaluate the PDFs and  $\mathcal{J}_{ij}(E, R, z, \mu)$  at the jet scale  $\mu_J = 2E \tan(R/2)$  and evolve the fragmentation function from  $2m_c$  to the scale  $\mu_J$  using the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equation,

$$\mu \frac{\partial}{\partial \mu} D_i(z, \mu) = \frac{\alpha_s(\mu)}{\pi} \sum_j \int_z^1 \frac{dy}{y} P_{i \rightarrow j}(z/y, \mu) D_j(y, \mu), \quad (4.30)$$

where  $P_{i \rightarrow j}(z/y, \mu)$  are the QCD splitting functions. We consider mixing between the charm quark and gluon splitting functions only for the  $^3S_1^{[1]}$  channel.<sup>9</sup> To leading order in  $\alpha_s$ , it can be shown that [62]

$$\frac{\mathcal{G}_i^\psi(E, R, z, \mu_J)}{2(2\pi)^3} \rightarrow D_{i \rightarrow \psi}(z, \mu_J) + \mathcal{O}(\alpha_s(\mu_J)). \quad (4.31)$$

Later in 4.3.3.2, we will also consider the 1-jet inclusive cross section. This is calculated by replacing the FJF in Eq. (4.28) with the jet function for a cone-type algorithm [138]. The FJFs are defined in Ref. [137] so that the sum over all possible fragmentations of a parton into hadrons equals the inclusive jet function.

$$J_i(E, R, \mu) = \frac{1}{2} \sum_h \int \frac{dz}{(2\pi)^3} z \mathcal{G}_i^h(E, R, z, \mu). \quad (4.32)$$

For further details about these calculations we refer the reader to Ref. [62]. Throughout this section we choose  $m_c = 1.4$  GeV and  $R = 0.4$ .

---

<sup>9</sup>The charm quark fragmentation into a  $J/\psi$  is dominated by the  $^3S_1^{[1]}$  channel because the color singlet and octet contributions start at same order in  $\alpha_s$  but the color octet channels are suppressed in the  $v$  expansion.

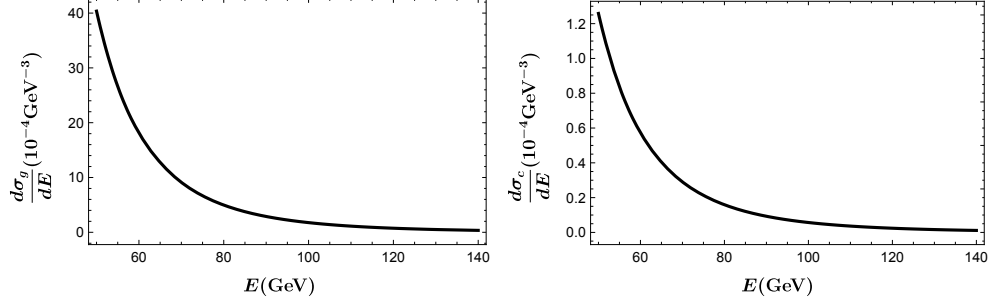


Figure 29: Cross sections for inclusive gluon and charm jets at the LHC. The center of mass energy is  $\sqrt{s} = 13$  TeV.

### 4.3.3 Discussion of the $J/\psi$ production mechanisms

In this subsection, we discuss the predictions for  $J/\psi$  production in jets using the LDMEs extracted by various groups and reveal some generic features that are independent of these extractions. The LDMEs we use in this section are summarized in Table 6. Refs. [2, 3] use a global fit to 194 data points from 26 data sets and predict significant polarization of the  $J/\psi$  in the high  $p_T$  region, which contradicts the measurements at the Tevatron [134] and the LHC [135, 136]. The extractions in Refs. [5, 6] focus on the high  $p_T$  region and attempt to solve the polarization puzzle.

**4.3.3.1 Normalized  $J/\psi$  production cross section** To discuss the dependence of  $J/\psi$  production on the associated jet energy, we use a normalized differential cross section defined as

$$\frac{d\tilde{\sigma}_i}{dEdz} \equiv \frac{d\sigma_i}{dEdz} \bigg/ \sum_i \int_{z_{min}}^{z_{max}} dz \frac{d\sigma_i}{dEdz}, \quad (4.33)$$

and

$$\frac{d\tilde{\sigma}}{dEdz} \equiv \sum_i \frac{d\tilde{\sigma}_i}{dEdz}, \quad (4.34)$$

where  $i$  denotes different  $J/\psi$  production channels (i.e., for the gluon initiated jets  $i \in \{^1S_0^{[8]}, ^3S_1^{[8]}, ^3P_J^{[8]}, ^3S_1^{[1]}\}$  and for the charm initiated jets  $i = ^3S_1^{[1]}$ ), and  $d\sigma_i/dEdz$  is defined in Eq. (4.28).

In Eq. (4.33),  $z_{min}$  ( $z_{max}$ ) should not be too close to 0 (1) where the factorization breaks down. The motivation for studying this normalized cross section is that we want to isolate the properties of quarkonium fragmentation in jets from the hard processes that generate the jet initiating partons. Fig. (29) shows the energy distributions of the hard process for gluon and charm jets at the LHC<sup>10</sup>. For all the figures in this section, we fix the center of mass energy to be  $\sqrt{s} = 13$  TeV.

Fig. (30) shows the comparison of the normalized (Eq. (4.33)) and unnormalized cross sections (Eq. (4.28)), where the LDMEs from Ref. [2, 3] are used with  $z_{min} = 0.3$  and  $z_{max} = 0.8$ . Corresponding plots for the LDMEs of Ref. [5] and Ref. [6] are shown in Appendix C.1 and C.2 respectively. We would like to emphasize the fact that both the unnormalized and normalized cross sections are directly measurable in experiments, although the normalized cross section has a better resolving power than the unnormalized cross section. In particular, the unnormalized cross section is a decreasing function of  $E$  for all the production channels due to the decreasing nature of the hard process, while the normalized cross section can be an increasing function for certain production channels due to the properties of their FJF's.

A measurement of the normalized cross section (Eq. (4.33)) for  $z > 0.5$ , can help identify both the dominant channel and the favored set of LDMEs. From Fig. (30), we can see that if  $d\tilde{\sigma}/dEdz$  turns out to be a decreasing function of the jet energy for  $z > 0.5$ , then the depolarizing  $^1S_0^{[8]}$  should be the dominant channel. We find this result to be true for LDME extractions of Ref. [5] as well (see Appendix C.1).

Note that in Fig. (30), the  $^3S_1^{[1]}$  channel makes the largest contribution and is also a decreasing function of the jet energy. So in principle a decreasing total normalized cross section could indicate the dominance of the  $^3S_1^{[1]}$  channel. However, NLO calculations in  $\alpha_s$  show that in the high  $p_T$  region, contribution from the  $^3S_1^{[1]}$  channel is very small and lies orders of magnitude below the data [15, 133, 5, 141, 142, 143, 144, 145].<sup>11</sup> Hence, if experiments find the normalized cross section to be a decreasing function of  $E$  for  $z > 0.5$ , then  $^1S_0^{[8]}$  must be the dominant channel.<sup>12</sup>

<sup>10</sup>We consider leading order partonic cross sections convoluted with PDF [139, 140], which includes the following processes:  $gg \rightarrow gg$ ,  $gq(\bar{q}) \rightarrow gq(\bar{q})$ ,  $q\bar{q} \rightarrow gg$ ,  $gg \rightarrow c\bar{c}$ ,  $gc(\bar{c}) \rightarrow gc(\bar{c})$ ,  $cc \rightarrow cc$ ,  $\bar{c}\bar{c} \rightarrow \bar{c}\bar{c}$ ,  $cq(\bar{q}) \rightarrow cq(\bar{q})$ ,  $\bar{c}q(\bar{q}) \rightarrow \bar{c}q(\bar{q})$ ,  $q\bar{q} \rightarrow c\bar{c}$ ,  $c\bar{c} \rightarrow c\bar{c}$ .

<sup>11</sup>We test this in Appendix C.3 by ignoring the  $^3S_1^{[1]}$  channel contribution to the normalization.

<sup>12</sup>We normalize the cross section by integrating from  $z = 0.3$  to 0.8 and for low values of  $z$ , the  $^3S_1^{[1]}$

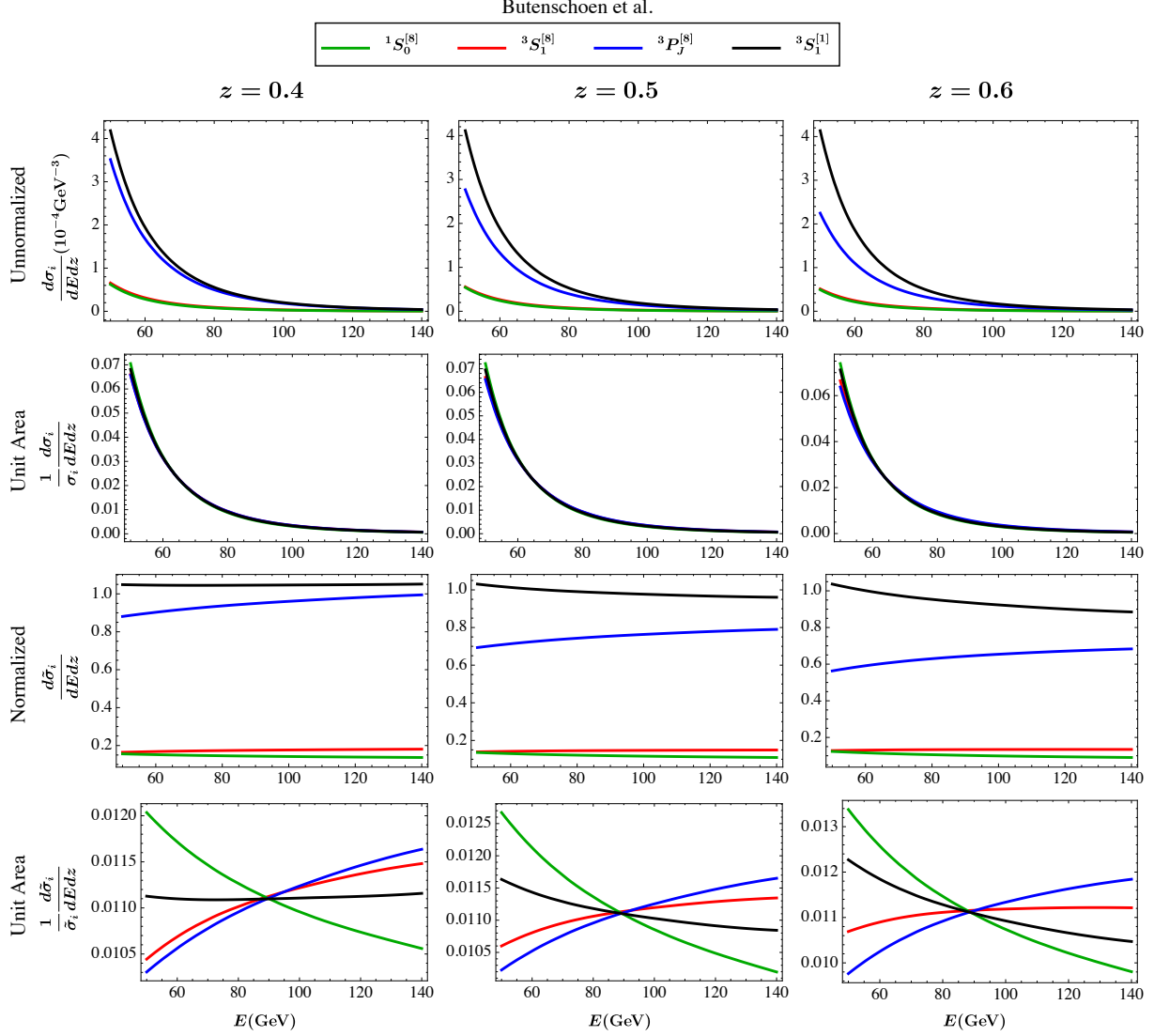


Figure 30: Cross sections for the different production channels at  $z = 0.4, 0.5$ , and  $0.6$  as a function of the jet energy. The first two rows show the unnormalized cross sections ( $d\sigma_i/dEdz$ ), with the second row showing plots normalized to unit area for a better visualization of the shapes, i.e., we multiply each curve of the first row by an appropriate constant to get the corresponding curve in the second row. Similar plots for the normalized cross section ( $d\tilde{\sigma}_i/dEdz$ ) are shown in the third and fourth row. The LDMEs are from Butenschoen et al.'s extractions [2].

In Fig. (31), we show the jet energy dependence of the total normalized cross sections (Eq. (4.34)) based on different LDME extractions. The error bands are purely due to the contribution can be significant. So even though the color singlet channel cannot dominate in the high  $p_T$  region, its contribution are not completely ignored in our analysis.

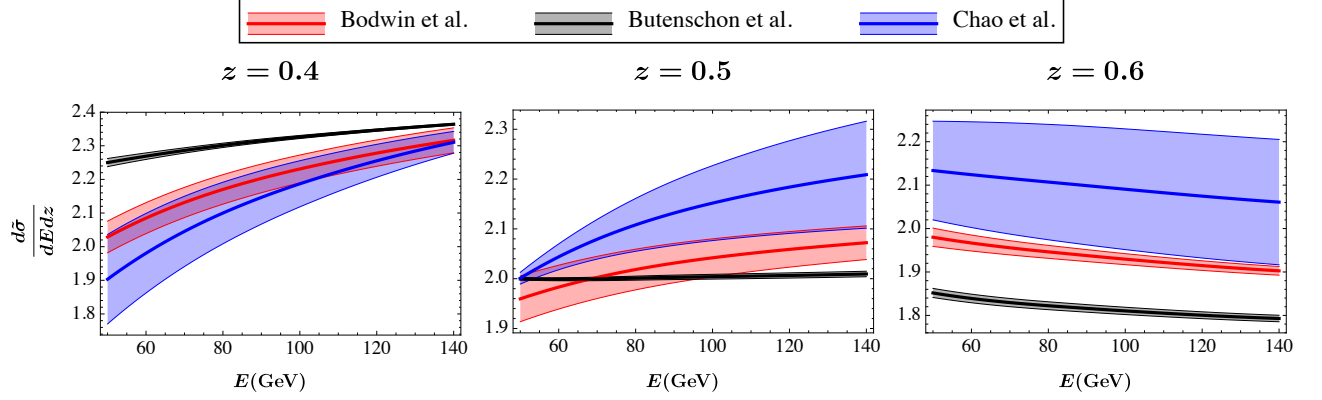


Figure 31: Total normalized cross section (i.e.  $d\hat{\sigma}/dEdz$  defined in Eq. (4.34)) with error bands. Red, black, and blue curves correspond to Bodwin et al. [5], Butenschoen et al. [2, 3], and Chao et al.’s [6] extractions, respectively.

LDME uncertainties, that is, we consider the uncertainty due to each LDME and sum by quadrature to obtain the total uncertainties<sup>13</sup>. It can be seen in Fig. (31) that as  $z$  goes from 0.4 to 0.6, the shapes change from an increasing function to a decreasing function. However since different extractions have distinct slopes, this observable has the potential power to test these extractions at the LHC. A different choice of  $(z_{min}, z_{max})$  does not change our arguments as we demonstrate in Appendix C.4.

**4.3.3.2 Normalization using 1-jet inclusive cross section** We now normalize the cross section in such a way that the denominator is independent of the LDMEs. This allows us to make a direct comparison of our results to those of Ref. [62]. The normalization is defined as

$$\frac{d\hat{\sigma}_i}{dEdz} \equiv \frac{d\sigma_i}{dEdz} \bigg/ \frac{d\sigma_J}{dE}, \quad (4.35)$$

and

$$\frac{d\hat{\sigma}}{dEdz} \equiv \sum_i \frac{d\hat{\sigma}_i}{dEdz}, \quad (4.36)$$

<sup>13</sup>To obtain the error bands corresponding to the extraction from Bodwin et al., we have used the error correlation matrix not shown in the original paper [121].

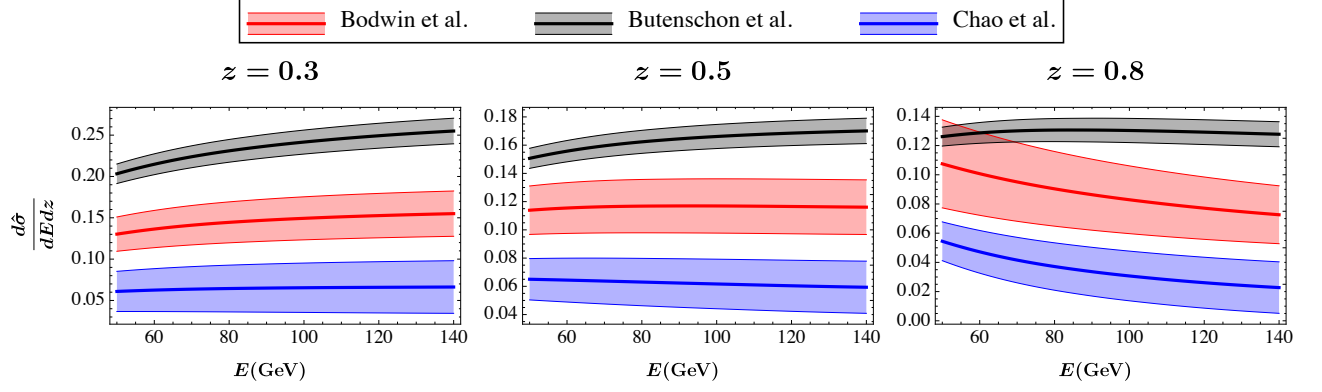


Figure 32: Total normalized cross section (i.e.  $d\hat{\sigma}/dEdz$  defined in Eq. (4.36)) with error bands. Red, black, and blue curves correspond to Bodwin et al. [5], Butenschon et al. [2, 3], and Chao et al.'s [6] extractions, respectively.

where  $d\sigma_i/dEdz$  is the same as that in Eq. (4.28) and  $d\sigma_J/dE$ <sup>14</sup> is the 1-jet inclusive cross section<sup>15</sup>. Note that the  $z$ -dependence of Eq. (4.35) comes only from the  $\mathcal{G}_i^{J/\psi}(E, R, z, \mu)$  in Eq. (4.28).

Fig. (32) shows the total  $J/\psi$  production cross section based on Eq. (4.36). The key feature of this plot is that the arguments given Ref. [62] based on the FJFs are also true for the cross section (see Fig. 6 in Ref. [62])<sup>16</sup>. Specifically, when  $z > 0.5$ , the shapes of the curves are very different for the extraction based on a global fit (black curves) and the other two based on fit to high  $p_T$  region (red and blue curves). Since the extractions from the global fit and high  $p_T$  fit give rise to different slopes for the  $J/\psi$  production cross section, one can test which set of the LDME extractions are preferred by measuring these slopes. Note that because our results are for the cross section, all the curves have positive values, in contrast to the gluon FJF for the LDMEs of Ref. [6] (shown in Fig. (6) of Ref. [62]) which became negative at large energies.

In Fig. (33), we plot the  $E$  dependence of the individual  $J/\psi$  production channels for

<sup>14</sup>This includes the contributions of gluon, light quarks, charm and bottom jets.

<sup>15</sup>The definition of Eq. (4.35) is essentially the same as the jet fragmentation function introduced in Ref. [124], except that we have integrated the jet pseudorapidity over the region  $|\eta_J| < 1.2$  for the denominator and numerator.

<sup>16</sup>To facilitate direct comparison of our Fig. (32) to Fig. (6) in Ref. [62], we make plots for  $z = 0.3, 0.5$  and  $0.8$ .



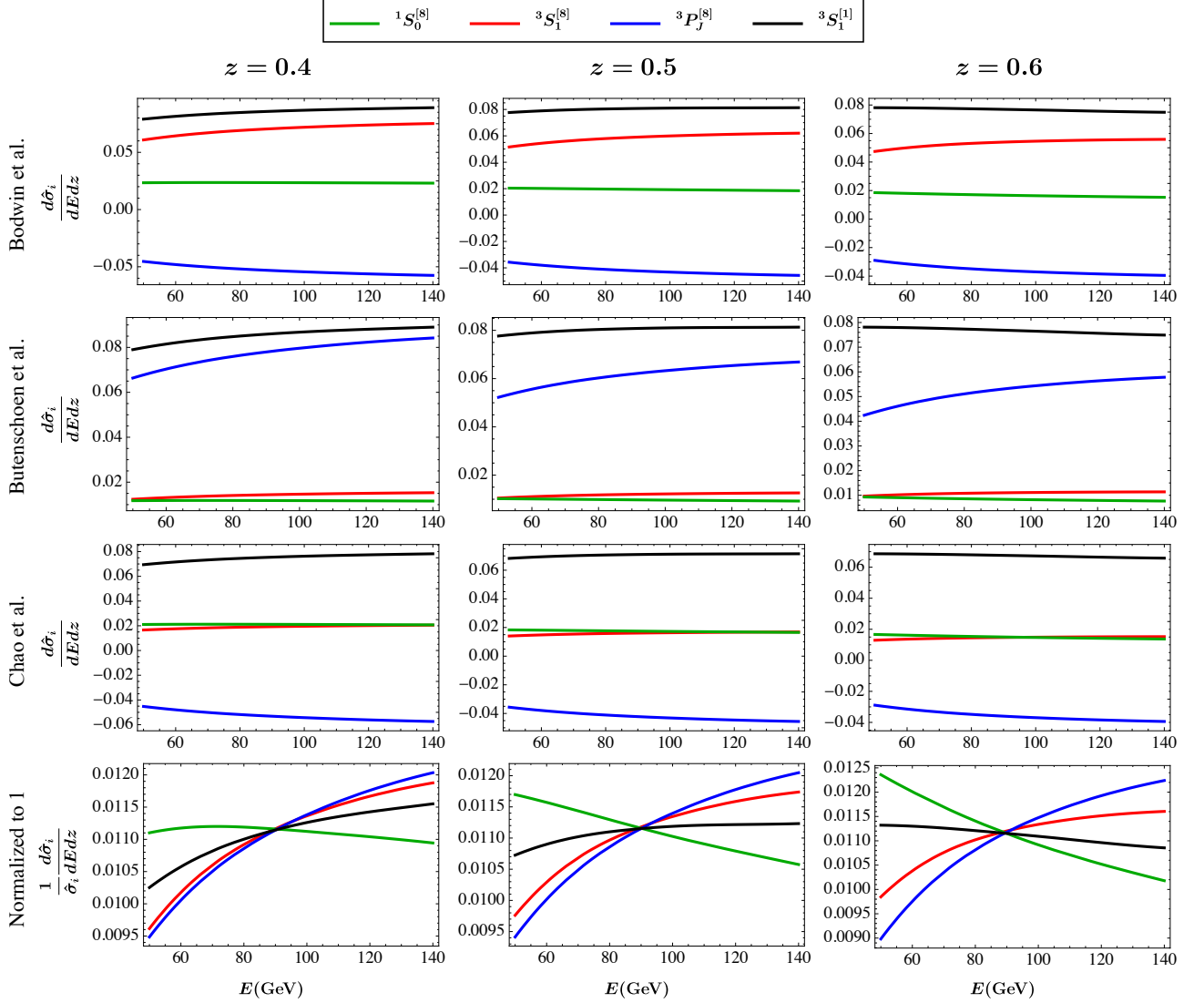


Figure 33: Comparisons of the production channels for various LDMEs using Eq. (4.35). Last row shows the plots normalized to unit area. This is indicated by  $1/\hat{\sigma}_i$  for the cross section label in the fourth row, which also cancels the LDME dependence of the numerator.

the different LDMEs using Eq. (4.35). We find that if the measurements of the observable defined in Eq. (4.35) results in a cross section which is a decreasing function of the jet energy for  $z > 0.5$ , then the  $^1S_0^{[8]}$  channel should have an anomalously large contribution to the  $J/\psi$  production. The fourth row in Fig. (33), with the curves normalized to unit area, clearly shows that only  $^1S_0^{[8]}$  channel is a decreasing function of jet energy for  $z > 0.5$ . Note that

in the fourth row of Fig. (33), the LDME dependence gets canceled due to normalization to unit area and so the prediction for  $^1S_0^{[8]}$  channel being dominant at high  $p_T$  is independent of any specific LDME extractions.<sup>17</sup>

To conclude this subsection, we mention a few things about the normalization conventions in Eq. (4.33) and Eq. (4.35). First of all, both the normalizations can be directly tested in experiments. Also since both the numerator and denominator of Eq. (4.33) depend on the LDMEs, they are statistically correlated and hence the width of error bands in Fig. (31) is reduced. However, Eq. (4.35) does not have such a correlation since the jet cross section used for the normalization is independent of the LDMEs. Indeed, if we look at Bodwin et al.’s extraction near  $z = 0.5$  and  $E = 100$  GeV, the ratio of the width of error band to the center value is  $\sim 4\%$  in Fig. (31) and  $\sim 30\%$  in Fig. (32). On the other hand, in both Fig. (31) and Fig. (32), the shapes of blue and red curves (high  $p_T$  fit) are in contrast to the black curve (global fit).

#### 4.3.4 Conclusion

In this section, we have looked at the total cross section for  $J/\psi$  production at the LHC by using the FJF approach. We make comparisons between the different NRQCD production channels for the  $J/\psi$ . We show that if for  $z > 0.5$  the normalized cross section is a decreasing function of jet energy, then the depolarizing  $^1S_0^{[8]}$  should be the dominant production channel at high  $p_T$ . We find this to be true for two sets of normalized cross sections. Our results confirm that the prediction made in Ref. [62] regarding the decreasing nature (with  $E$ ) of the FJF for  $^1S_0^{[8]}$  channel, does not change by inclusion of the hard scattering effects. Using our normalized cross sections, one can also test which set of the LDME extractions are favored.

---

<sup>17</sup>In Fig. (30), both the  $^3S_1^{[1]}$  and  $^1S_0^{[8]}$  were found to be decreasing functions of  $E$  and so this observable provides an additional tool to distinguish between these two channels.

## 5.0 SUMMARY

We study the physics of jets and quarkonium production using the QCD effective theories SCET and NRQCD. Specifically we did work related to the following two topics.

### Jet Physics:

In Refs. [31, 32], we introduced a function called fragmentation function to a jet (FFJ) to describe inclusive jet production from a parton and studied FFJs in different phase space and momentum regions. Specifically, in Ref. [31], we investigated the situation where the jet radius was small. Jets with small radii could appear in different scenarios. Nothing stops us to measure jets with small radii except for the limit of experimental detectors. For instance, at the LHC, QCD jets with radius as small as  $r = 0.1$  could be reliably measured. For jets with small radius  $r$ , large  $\ln(r)$  appears that needs to be resummed. As another example, when jet substructures are considered,  $\ln(R/r)$  would naturally appear, where  $R$  and  $r$  correspond to the fat- and sub-jet radius, respectively. Another limit we investigated in Ref. [32] was the large  $z$  limit, where  $z$  is defined to be the fraction of energy carried by the jet from the mother parton. This is the limit where QCD dynamics gradually becomes non-perturbative and large  $\ln(1 - z)$  would appear. We found that the emergence of this  $\ln(1 - z)$  was due to the gluon radiations that were both collinear and soft, and could be described by the collinear-soft mode under the SCET framework. We formulated factorization theorems and used renormalization group techniques to deal with these types of logarithms.

### Quarkonium Production in Jets:

In Ref. [87], we studied electron-positron collisions and investigated  $B$  meson and  $J/\psi$  production in jets. We analyzed the energy distribution of a  $B$  meson or  $J/\psi$  in jets with different

jet shapes, both analytically and using Monte Carlo simulations with PYTHIA and Herwig.  $B$  meson production in jets served as a test of our theoretical framework, since  $B$  meson fragmentation functions (based on empirical models) were experimentally well measured, and these models could be chosen in PYTHIA for the hadronization process. We found that our analytic calculations and PYTHIA simulations agreed very well in this case. However, when we applied the same framework to quarkonium production, analytic calculations and the default PYTHIA simulations gave very different results. After investigating PYTHIA, we found its treatment of quarkonium fragmentation to be too primitive. We attempted to modify PYTHIA to incorporate NRQCD fragmentation functions with the parton shower and found that the modified PYTHIA agreed with our analytic calculations.

The LHCb collaboration recently measured  $J/\psi$  production in jets. They also found that the default PYTHIA simulations had similar features as our analysis in Ref. [87] and were at odds with the collected data. We generalized our analysis of electron-positron collisions to LHC events in Ref. [88] and found better agreements with data. Further more, according to our analysis, long distance matrix elements (LDMEs) extracted from high  $p_T$  regions gave better description of the LHCb data than those from a global fit (which used a large set of data, including those at low  $p_T$  regions). In Ref. [89], we were still considering  $J/\psi$  production in jets at the LHC, but looking at energy distributions of those jets that have a  $J/\psi$  produced in them. We showed that these observables could have the potential power to discriminate different quarkonium production mechanisms and test which LDME extractions are preferable.

## APPENDIX A

### A.1 HADRON FRAGMENTATION INSIDE A JET

We can describe the HFF inside a jet,  $D_{H/J_k}(z)$ , similar to sJFF. The unnormalized HFF inside a jet can be expressed as

$$\begin{aligned} \tilde{D}_{H/J_q}(z, \mu) &= \frac{z^{D-3}}{2N_c} \sum_{X \in j} \text{Tr} \langle 0 | \delta\left(\frac{p_H^+}{z} - \mathcal{P}_+\right) \frac{\not{n}}{2} \Psi_n | H X \in J(p_J^+, R) \rangle \\ &\quad \times \langle H X \in J(p_J^+, R) | \bar{\Psi}_n | 0 \rangle. \end{aligned} \quad (\text{A.1})$$

Here we described the hardron fragmentation from the quark jet in the hadron frame ( $\mathbf{p}_H^\perp = 0$ ), and the momentum of the mother parton is given by  $p_J$ , hence  $zp_J^+ = p_H^+$ .

Although the fragmentation function is a nonperturbative observable, it is important to understand its renormalization behavior computing the higher order corrections at the parton level separating IR divergences. At LO in  $\alpha_s$ , the fragmentation function from quark jet to quark is given by  $D_{q/J_q}^{(0)}(z) = \delta(1-z)$ . At NLO in  $\alpha_s$ , the virtual correction, including zero-bin subtraction, is

$$D_V = \frac{\alpha_s C_F}{\pi} \left( \frac{1}{\varepsilon_{\text{UV}}} - \frac{1}{\varepsilon_{\text{IR}}} \right) \left( \frac{1}{\varepsilon_{\text{UV}}} + \ln \frac{\mu}{p_J^+} + 1 \right) \delta(1-z). \quad (\text{A.2})$$

The Feynman diagrams for real gluon emissions are shown in Fig. 34, and only diagram Fig. 34-(a) has a nonvanishing zero-bin contribution. Thus the amplitude for Fig. 34-(a) is written as

$$D_R^{(a)} = \tilde{D}_R^{(a)} - D_{R,0}^{(a)}, \quad (\text{A.3})$$

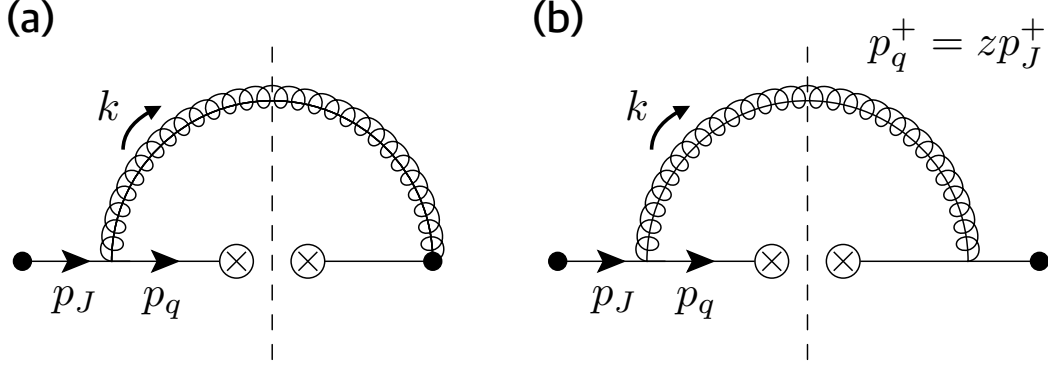


Figure 34: Feynman diagrams of real gluon emissions for quark fragmentation inside a jet at NLO. The gluon in the final state is also inside a jet. Diagram (a) has its Hermitian conjugate contribution.

where  $\tilde{D}$  is the naive collinear contribution and  $D_0$  is the zero-bin contribution.  $\tilde{D}$  is given by

$$\begin{aligned}\tilde{D}_R^{(a)}(z) &= \frac{\alpha_s C_F}{2\pi} \frac{(\mu^2 e^\gamma)^\varepsilon}{\Gamma(1-\varepsilon)} \int_0^{\Lambda_{\text{alg}}^2 = p_J^{+2} t^2 z(1-z)} \frac{dM^2}{(M^2)^{1+\varepsilon}} z^{1-\varepsilon} (1-z)^{-1-\varepsilon} \\ &= \tilde{I}_R^{(a)} \delta(1-z) + \left[ D_R^{(a)}(z) \right]_+, \end{aligned} \quad (\text{A.4})$$

where  $\tilde{I}_R^{(a)}$  is

$$\begin{aligned}\tilde{I}_R^{(a)} &= \int_0^1 dz \tilde{D}_R^{(a)}(z) \\ &= \frac{\alpha_s C_F}{2\pi} \left[ \frac{1}{2\varepsilon_{\text{IR}}^2} + \frac{1}{\varepsilon_{\text{IR}}} \left( 1 + \ln \frac{\mu}{p_J^+ t} \right) + 4 - \frac{3\pi^2}{8} + \ln \frac{\mu}{p_J^{+2} t^2} + \frac{1}{4} \ln^2 \frac{\mu}{p_J^{+2} t^2} \right]. \end{aligned} \quad (\text{A.5})$$

For the zero-bin contribution, the radiated gluon becomes soft and hence the  $z$ -dependence can be fixed as  $\delta(1-z)$ , giving

$$\begin{aligned}D_{R,0}^{(a)}(z) &= \frac{\alpha_s C_F}{2\pi} \frac{(\mu^2 e^\gamma)^\varepsilon}{\Gamma(1-\varepsilon)} \delta(1-z) \int_0^\infty dk_+ k_+^{-1-\varepsilon} \int_0^{t^2 k_+} k_-^{-1-\varepsilon} \\ &= \frac{\alpha_s C_F}{2\pi} \left[ \frac{1}{2} \left( \frac{1}{\varepsilon_{\text{UV}}} - \frac{1}{\varepsilon_{\text{IR}}} \right)^2 + \left( \frac{1}{\varepsilon_{\text{UV}}} - \frac{1}{\varepsilon_{\text{IR}}} \right) \ln t \right] \delta(1-z), \end{aligned} \quad (\text{A.6})$$

where the phase space constraint by the jet algorithm gives  $t^2 > k_-/k_+$  from Eqs. (3.9) and (3.10), and the jet mass is approximated as  $M^2 \sim p_J^+ k_-$ .

Similar to Eq. (A.4), the contribution of diagram Fig 34-(b) is

$$\begin{aligned} D_R^{(b)}(z) &= \frac{\alpha_s C_F}{2\pi} \frac{(\mu^2 e^\gamma)^\varepsilon}{\Gamma(1-\varepsilon)} (1-\varepsilon) \int_0^{\Lambda_{\text{alg}}^2 = p_J^{+2} t^2 z(1-z)} \frac{dM^2}{(M^2)^{1+\varepsilon}} z^{-\varepsilon} (1-z)^{1-\varepsilon} \\ &= I_R^{(b)} \delta(1-z) + \left[ M_R^{(b)}(z) \right]_+, \end{aligned} \quad (\text{A.7})$$

where the integrated part  $I_R^{(b)}$  is

$$I_R^{(b)} = -\frac{\alpha_s C_F}{2\pi} \left[ \frac{1}{2\varepsilon_{\text{IR}}} + \ln \frac{\mu}{p_J^+ t} + \frac{3}{2} \right]. \quad (\text{A.8})$$

Therefore combining Eqs. (A.2), (A.5), (A.6), and (A.8), we can obtain the part proportional to  $\delta(1-z)$ . This result should be equal to Eq. (3.14), i.e., the integrated jet function at NLO for  $\theta < R'$ . This is confirmed by

$$\begin{aligned} I_{q/q}^{\theta < R} \delta(1-z) &= \mathcal{J}_q^{(1)}(\mu; E_J R') \delta(1-z) \\ &= D_V + 2(\tilde{I}_R^{(a)} \delta(1-z) - M_{R,0}^{(a)}(z)) + I_R^{(b)} \delta(1-z) + (Z_\xi^{(1)} + R_\xi^{(1)}) \delta(1-z) \\ &= \delta(1-z) \frac{\alpha_s C_F}{2\pi} \left[ \frac{1}{\varepsilon_{\text{UV}}^2} + \frac{1}{\varepsilon_{\text{UV}}} \left( \frac{3}{2} + \ln \frac{\mu^2}{p_J^{+2} t^2} \right) \right. \\ &\quad \left. + \frac{3}{2} \ln \frac{\mu^2}{p_J^{+2} t^2} + \frac{1}{2} \ln^2 \frac{\mu^2}{p_J^{+2} t^2} + \frac{13}{2} - \frac{3\pi^2}{4} \right], \end{aligned} \quad (\text{A.9})$$

where  $Z_\xi$  is the collinear quark field strength renormalization and  $R_\xi$  is its residue. At one loop they are given by

$$Z_\xi^{(1)} = -\frac{\alpha_s C_F}{4\pi} \frac{1}{\varepsilon_{\text{UV}}}, \quad R_\xi^{(1)} = \frac{\alpha_s C_F}{4\pi} \frac{1}{\varepsilon_{\text{IR}}}. \quad (\text{A.10})$$

The remaining distribution parts in Eqs. (A.4) and (A.7) are

$$\begin{aligned} \left[ D_R(z) \right]_+ &= \left[ 2D_R^{(a)}(z) + D_R^{(b)}(z) \right]_+ \\ &= -\frac{\alpha_s C_F}{2\pi} \left[ \frac{1+z^2}{1-z} \left( \frac{1}{\varepsilon_{\text{IR}}} + \ln \frac{\mu^2}{p_J^{+2} t^2} - 2 \ln z(1-z) \right) - (1-z) \right]_+ \end{aligned} \quad (\text{A.11})$$

$$\begin{aligned} &= -\frac{\alpha_s C_F}{2\pi} \left\{ \delta(1-z) \left[ \frac{3}{2} \left( \frac{1}{\varepsilon_{\text{IR}}} + \ln \frac{\mu^2}{p_J^{+2} t^2} \right) + \frac{13}{2} - \frac{2\pi^2}{3} \right] \right. \\ &\quad \left. + (1+z^2) \left[ \frac{1}{(1-z)_+} \left( \frac{1}{\varepsilon_{\text{IR}}} + \ln \frac{\mu^2}{p_J^{+2} t^2} - 2 \ln z \right) - 2 \left( \frac{\ln(1-z)}{1-z} \right)_+ \right] - (1-z) \right\}. \end{aligned} \quad (\text{A.12})$$

Finally, combining Eqs. (A.9) and (A.12), we obtain the unnormalized HFF inside a jet up to NLO,

$$\begin{aligned}\tilde{D}_{q/J_q}(z, \mu; E_J R') &= \mathcal{J}_q(\mu; E_J R') \delta(1-z) + \left[ D_R(z) \right]_+ \\ &= \delta(1-z) \left\{ 1 + \frac{\alpha_s C_F}{2\pi} \left[ \frac{1}{\varepsilon_{\text{UV}}^2} + \frac{1}{\varepsilon_{\text{UV}}} \ln \frac{\mu^2}{p_J^{+2} t^2} + \frac{3}{2} \left( \frac{1}{\varepsilon_{\text{UV}}} - \frac{1}{\varepsilon_{\text{IR}}} \right) + \frac{1}{2} \ln^2 \frac{\mu^2}{p_J^{+2} t^2} - \frac{\pi^2}{12} \right] \right\} \\ &\quad - \frac{\alpha_s C_F}{2\pi} \left\{ (1+z^2) \left[ \frac{1}{(1-z)_+} \left( \frac{1}{\varepsilon_{\text{IR}}} + \ln \frac{\mu^2}{p_J^{+2} t^2} - 2 \ln z \right) - 2 \left( \frac{\ln(1-z)}{1-z} \right)_+ \right] - (1-z) \right\}.\end{aligned}\quad (\text{A.13})$$

The normalized HFF inside a jet is obtained by dividing by  $\mathcal{J}_q(\mu; E_J R')$ ,

$$\begin{aligned}D_{q/J_q}(z) &= \frac{\tilde{D}_{q/J_q}(z; E_J R')}{\mathcal{J}_q(\mu; E_J R')} \\ &= \delta(1-z) - \frac{\alpha_s}{2\pi} \left\{ P_{qq}(z) \left( \frac{1}{\varepsilon_{\text{IR}}} + \ln \frac{\mu^2}{p_J^{+2} t^2} \right) + C_F \left[ \delta(1-z) \left( \frac{13}{2} - \frac{2\pi^2}{3} \right) - (1-z) \right. \right. \\ &\quad \left. \left. - 2(1+z^2) \left( \frac{\ln z}{(1-z)_+} + \left( \frac{\ln(1-z)}{1-z} \right)_+ \right) \right] \right\}.\end{aligned}\quad (\text{A.14})$$

In a similar way we can compute the other HFFs inside a jet. Their NLO results are

$$D_{g/J_q}(z) = -\frac{\alpha_s}{2\pi} \left[ P_{gq}(z) \left( \frac{1}{\varepsilon_{\text{IR}}} + \ln \frac{\mu^2}{p_J^{+2} t^2} - 2 \ln z(1-z) \right) - z C_F \right], \quad (\text{A.15})$$

$$\begin{aligned}D_{g/J_g}(z) &= \delta(1-z) - \frac{\alpha_s}{2\pi} \left\{ P_{gg}(z) \left( \frac{1}{\varepsilon_{\text{IR}}} + \ln \frac{\mu^2}{p_J^{+2} t^2} \right) + N_c \left[ \delta(1-z) \left( \frac{67}{9} - \frac{23n_f}{18N_c} - \frac{2\pi^2}{3} \right) \right. \right. \\ &\quad \left. \left. - 4 \left[ \frac{z \ln z}{(1-z)_+} + z \left( \frac{\ln(1-z)}{1-z} \right)_+ + \ln[z(1-z)] \left( \frac{1-z}{z} + z(1-z) \right) \right] \right] \right\},\end{aligned}\quad (\text{A.16})$$

$$D_{q/J_g}(z) = -\frac{\alpha_s}{2\pi} \left[ P_{qg}(z) \left( \frac{1}{\varepsilon_{\text{IR}}} + \ln \frac{\mu^2}{p_J^{+2} t^2} - 2 \ln[z(1-z)] \right) - z(1-z) \right]. \quad (\text{A.17})$$

At much lower energy scale,  $\mu \ll p_J^+ t$ , the fragmenting process cannot resolve the scale  $p_J^+ t$ . Hence the scale  $p_J^+ t$  can be identified as an UV scale. In this case the fragmenting process can be described by the standard FF without the phase space restriction. Therefore, similar to the subjet case shown in Eq. (3.70), the FF inside a jet is in general factorized as follows [55]:

$$D_{l/J_k}(z, \mu; E_J R') = \int_z^1 \frac{dx}{x} K_{m/k}(z/x, \mu; E_J R') D_{l/m}(x, \mu), \quad (\text{A.18})$$



where  $k$ ,  $l$ , and  $m$  represents the quark flavors and gluon, and  $m$  is the dummy index. If we consider the hadron FF, we have

$$D_{H/J_k}(z, \mu; E_J R') = \int_z^1 \frac{dx}{x} K_{m/k}(z/x, \mu; E_J R') D_{H/m}(x, \mu). \quad (\text{A.19})$$

Here  $K_{m/k}$  are the perturbative kernels with a typical energy scale  $p_J^+ t \sim E_J R'$ . They are obtained from the matching between two FFs. Because  $K_{m/k}$  is irrelevant to the lower energy scale dynamics, they are universally given when we consider a fragmentation process inside a jet.

Under dimensional regularization, the bare result of NLO corrections to the standard FF at parton level is

$$D_{l/m}^{(1)}(z) = \frac{\alpha_s}{2\pi} P_{lm}(z) \left( \frac{1}{\varepsilon_{\text{UV}}} - \frac{1}{\varepsilon_{\text{IR}}} \right), \quad (\text{A.20})$$

where  $P_{lm}$  are DGLAP splitting kernels. Comparing the NLO results of the HFF inside a jet and Eq. (A.20), we can easily check that the kernels in Eqs. (A.18) and (A.19) are the same as ones for the subjet case in Subsection 3.1.4.

## APPENDIX B

### B.1 RENORMALIZATION GROUP AND RESUMMATION

#### B.1.1 Evolution of Measured and Unmeasured Functions

The RGEs satisfied by the elements of the factorization theorem are separated into two categories; terms that do depend on the variable  $\tau_a$  and terms that do not. The latter satisfy the following RGE

$$\mu \frac{d}{d\mu} f(\mu) = \gamma_f(\mu) f(\mu), \quad (\text{B.1})$$

where  $\gamma_F(\mu)$  is the anomalous dimension

$$\gamma_F(\mu) = -\frac{1}{Z_F(\mu)} \mu \frac{d}{d\mu} Z_F(\mu) = \Gamma_F(\alpha_s) \ln \left( \frac{\mu^2}{m_F^2} \right) + \gamma_F(\alpha_s), \quad (\text{B.2})$$

where  $m_F$  is related to the characteristic scale for the particular function, and  $Z_F(\mu)$  is the renormalization function for  $F(\mu)$ . The coefficient  $\Gamma_F(\alpha_s)$  is proportional to the cusp anomalous dimension,  $\Gamma_{\text{cusp}}(\alpha_s)$ , which can be expanded in  $\alpha_s$

$$\Gamma_{\text{cusp}}(\alpha_s) = \sum_{n=0}^{\infty} \left( \frac{\alpha_s}{4\pi} \right)^{1+n} \Gamma_c^n, \quad (\text{B.3})$$

and  $\Gamma_F = (\Gamma_F^0/\Gamma_c^0) \Gamma_{\text{cusp}}$ . The non-cusp part,  $\gamma_F(\alpha_s)$ , has a similar expansion

$$\gamma_F(\alpha_s) = \sum_{i=0}^{\infty} \left( \frac{\alpha_s}{4\pi} \right)^{1+i} \gamma_F^i. \quad (\text{B.4})$$

The solution to RGE is given by

$$F(\mu) = \exp(K_F(\mu, \mu_0)) \left( \frac{\mu_0}{m_F} \right)^{\omega_F(\mu, \mu_0)} F(\mu_0), \quad (\text{B.5})$$

where the exponents  $K_F$  and  $\omega_F$  are given in terms of the anomalous dimension,

$$K_F(\mu, \mu_0) = 2 \int_{\alpha(\mu)}^{\alpha(\mu_0)} \frac{d\alpha}{\beta(\alpha)} \Gamma_F(\alpha) \int_{\alpha(\mu_0)}^{\alpha} \frac{d\alpha'}{\beta(\alpha')} + \int_{\alpha(\mu)}^{\alpha(\mu_0)} \frac{d\alpha}{\beta(\alpha)} \gamma_F(\alpha), \quad (\text{B.6})$$

$$\omega_F(\mu, \mu_0) = 2 \int_{\alpha(\mu)}^{\alpha(\mu_0)} \frac{d\alpha}{\beta(\alpha)} \Gamma_F(\alpha), \quad (\text{B.7})$$

and for up to NLL and NLL' accuracy are given by

$$K_F(\mu, \mu_0) = -\frac{\gamma_F^0}{2\beta_0} \ln r - \frac{2\pi\Gamma_F^0}{(\beta_0)^2} \left[ \frac{r-1+r\ln r}{\alpha_s(\mu)} + \left( \frac{\Gamma_c^1}{\Gamma_c^0} - \frac{\beta_1}{\beta_0} \right) \frac{1-r+\ln r}{4\pi} + \frac{\beta_1}{8\pi\beta_0} \ln^2 r \right], \quad (\text{B.8})$$

$$\omega_F(\mu, \mu_0) = -\frac{\Gamma_F^0}{j_F\beta_0} \left[ \ln r + \left( \frac{\Gamma_c^1}{\Gamma_c^0} - \frac{\beta_1}{\beta_0} \right) \frac{\alpha_s(\mu_0)}{4\pi} (r-1) \right], \quad (\text{B.9})$$

where  $r = \alpha(\mu)/\alpha(\mu_0)$  and  $\beta_n$  are the coefficients of the QCD  $\beta$ -function,

$$\beta(\alpha_s) = \mu \frac{d\alpha_s}{d\mu} = -2\alpha_s \sum_{n=0}^{\infty} \left( \frac{\alpha_s}{4\pi} \right)^{1+n} \beta_n. \quad (\text{B.10})$$

The RGEs for functions that depend on the variable  $\tau_a$  are of the form

$$\mu \frac{d}{d\mu} F(\tau_a, \mu) = \left[ \gamma_F(\mu) \otimes F(\mu) \right] (\tau_a), \quad (\text{B.11})$$

where

$$\begin{aligned} \gamma_F(\tau_a, \mu) &= - \left[ Z_F^{-1}(\mu) \otimes \mu \frac{d}{d\mu} Z_F(\mu) \right] (\tau_a) \\ &= \Gamma_F(\alpha_s) \left( \ln \frac{\mu^2}{m_F^2} - \frac{2}{j_F} \left( \frac{\Theta(\tau_a)}{\tau_a} \right)_+ \right) + \gamma_F(\alpha_s) \delta(\tau_a), \end{aligned} \quad (\text{B.12})$$

and the solution to this equation is given by

$$F(\tau_a, \mu) = \exp(K_F + \gamma_E \omega_F) \frac{1}{\Gamma(-\omega_F)} \left( \frac{\mu_0}{m_F} \right)^{j_F \omega_F} \left[ \left( \frac{\Theta(\tau_a)}{(\tau_a)^{1+\omega_F}} \right)_+ \otimes F(\tau_a, \mu_0) \right]. \quad (\text{B.13})$$

### B.1.2 Plus-distribution identities

We begin with the equation

$$\int d\tau'' \left[ \frac{\Theta(\tau - \tau'')}{(\tau - \tau'')^{1+\omega_1}} \right]_+ \left[ \frac{\Theta(\tau'' - \tau')}{(\tau'' - \tau')^{1+\omega_2}} \right]_+ = \frac{\Gamma(-\omega_1)\Gamma(-\omega_2)}{\Gamma(-\omega_1 - \omega_2)} \left[ \frac{\Theta(\tau - \tau')}{(\tau - \tau')^{1+\omega_1+\omega_2}} \right]_+, \quad (\text{B.14})$$

which can be easily proven using Laplace transforms and the defining equation of the plus distribution,

$$[f(\tau)]_+ \equiv \lim_{\beta \rightarrow 0} \frac{d}{d\tau} [\theta(\tau - \beta)F(\tau)] , \quad (\text{B.15})$$

where  $F(\tau)$  is defined as

$$F(\tau) \equiv \int_1^\tau d\tau' f(\tau') , \quad (\text{B.16})$$

which yields

$$\mathcal{L} \left\{ \left( \frac{1}{\tau^{1+\omega}} \right)_+ \right\} = s^\omega \Gamma(-\omega) . \quad (\text{B.17})$$

The following equations can be derived by setting  $\tau' \rightarrow 0$  in Eq. (B.14), expanding in  $\omega_2$  both sides and matching powers:

$$\int d\tau' \left[ \frac{\Theta(\tau - \tau')}{(\tau - \tau')^{1+\omega}} \right]_+ \delta(\tau') = \left[ \frac{\Theta(\tau)}{\tau^{1+\omega}} \right]_+, \quad (\text{B.18})$$

$$\int d\tau' \left[ \frac{\Theta(\tau - \tau')}{(\tau - \tau')^{1+\omega}} \right]_+ \left[ \frac{\Theta(\tau')}{\tau'} \right]_+ = \left[ \frac{\Theta(\tau)}{\tau^{1+\omega}} \right]_+ (\ln \tau - H(-1 - \omega)) ,$$

$$\int d\tau' \left[ \frac{\Theta(\tau - \tau')}{(\tau - \tau')^{1+\omega}} \right]_+ \left[ \frac{\Theta(\tau') \ln \tau'}{\tau'} \right]_+ = \left[ \frac{\Theta(\tau)}{\tau^{1+\omega}} \right]_+ \frac{(\ln \tau - H(-1 - \omega))^2 + \pi^2/2 - \psi^{(1)}(-\omega)}{2} ,$$

where we used [46]

$$\left[ \frac{\Theta(\tau)}{\tau^{1+\omega}} \right]_+ = -\frac{1}{\omega} \delta(\tau) + \sum_{n=0}^{\infty} (-\omega)^n \left[ \frac{\Theta(\tau) \ln^n \tau}{\tau} \right]_+ . \quad (\text{B.19})$$

### B.1.3 Reorganization of logarithms of $(1 - z)$

The convolutions in the variable  $z$  need to be performed numerically since they involve the evolved FFs, which are evaluated by solving the DGLAP equation using Mellin transformations. For this reason we expand the plus-distributions using the following relations

$$\int_z^1 \frac{dx}{x} \left( \frac{1}{1-x} \right)_+ f\left(\frac{z}{x}\right) = \int_z^1 dx \frac{1}{1-x} \left( \frac{1}{x} f\left(\frac{z}{x}\right) - f(z) \right) + f(z) \ln(1-z), \quad (\text{B.20})$$

$$\int_z^1 \frac{dx}{x} \left( \frac{\ln(1-x)}{1-x} \right)_+ f\left(\frac{z}{x}\right) = \int_z^1 dx \frac{\ln(1-x)}{1-x} \left( \frac{1}{x} f\left(\frac{z}{x}\right) - f(z) \right) + f(z) \frac{1}{2} \ln^2(1-z). \quad (\text{B.21})$$

Thus for every function  $D(z)$  the convolution with  $f_{\mathcal{J}}^{ij}(\tau, z, \mu)$  gives

$$\begin{aligned} \frac{1}{T_{ij}} \frac{2\pi}{\alpha_s(\mu)} f_{\mathcal{J}}^{ij}(\tau, z, \mu) \bullet D(z) &= \delta_{ij} f_1(\tau, z, \mu) D(z) - \int_z^1 dx f_2(\tau, x, \mu) \left( \frac{\bar{P}_{ji}(x)}{x} \circ D\left(\frac{z}{x}\right) \right) \\ &+ \int_z^1 dx \left[ c_{ij}(x) - \frac{1}{1-a/2} \ln \left( 1 + \left( \frac{1-x}{x} \right)^{1-a} \right) \frac{\bar{P}_{ji}(x)}{x} \right] \circ D\left(\frac{z}{x}\right), \end{aligned} \quad (\text{B.22})$$

where

$$f_2(\tau, z, \mu) = 2 \ln \left( \frac{\mu}{\mu_J(\tau, z)} \right) + \frac{1}{1-a/2} H(-1-\Omega), \quad (\text{B.23})$$

with

$$\mu_J(\tau, z) = \omega \tau^{1/(2-a)} (1-z)^{(1-a)/(2-a)},$$

$$f_1(\tau, z, \mu) = \frac{1-a/2}{1-a} \left( f_2(\tau, z, \mu) \right)^2 + \frac{a(1-a/4)}{(1-a)(1-a/2)} \frac{\pi^2}{6} - \frac{1}{(1-a)(1-a/2)} \psi^{(1)}(-\Omega), \quad (\text{B.24})$$

$$\begin{aligned} c_{qq}(z) &= \frac{1-z}{z}, \\ c_{gg}(z) &= 0, \\ c_{gq}(z) &= 2(1-z), \\ c_{qg}(z) &= 1, \end{aligned}$$

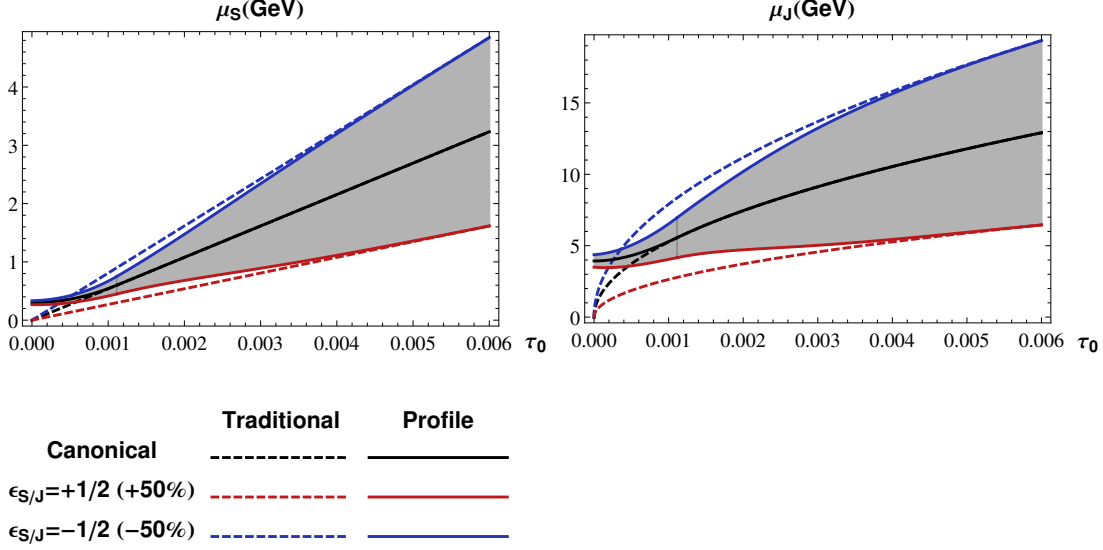


Figure 35: Profile functions for  $\mu_S^{PF}(\tau_0)$  and  $\mu_J^{PF}(\tau_0)$ , the  $\tau_0$ -dependent renormalization scales that we use in the scale variations of our measured soft function and measured jet function. Also shown are traditional scale variations done by varying  $\mu$  by  $\pm 50\%$ .

and

$$\begin{aligned}
 f(x) \circ g(x) &= f(x)g(x), \\
 [f(x)(h(x))_+] \circ g(x) &= h(x)[f(x)g(x) - f(1)g(1)].
 \end{aligned}$$

#### B.1.4 Profile Functions

Here, we write down the profile functions used to perform scale variations for our measured soft and measured jet functions. We use profile functions to introduce a  $\tau_a$ -dependent scale variation that freezes at the characteristic scale for high values of  $\tau_a$  where the factorization theorem breaks down and at a fixed scale for small values of  $\tau_a$  where we reach the non-perturbative regime. The profile function for the measured soft function,  $\mu_S^{PF}(\tau_0)$ , and the profile function for the measured jet function,  $\mu_J^{PF}(\tau_0)$ , are plotted in Fig. 35 (for the case

$a = 0$ ). The analytic formulae for these functions are

$$\mu_S^{PF}(\tau_a) = \left[ 1 + \epsilon_S \frac{g(\tau_a)}{g(1)} \right] \times \begin{cases} \mu_{min} + \alpha \tau_a^\beta & 0 < \tau_a < \tau_{min} \\ \omega \tau_a / r^{(1-a)} & \tau_{min} \leq \tau_a \end{cases}, \quad (\text{B.25})$$

$$\mu_J^{PF}(\tau_a) = \left[ 1 + \epsilon_J \frac{g(\tau_a)}{g(1)} \right] \times \begin{cases} (\omega r)^{(1-a)/(2-a)} (\mu_{min} + \alpha \tau_a^\beta)^{1/(2-a)} & 0 < \tau_a < \tau_{min} \\ \omega \tau_a^{1/(2-a)} & \tau_{min} \leq \tau_a \end{cases},$$

where we have defined

$$g(\tau) = \frac{1}{\exp \left( 1.26(\tau_{min} - \tau)/\tau_{min} \right) + 1}, \quad (\text{B.26})$$

and where  $\alpha$  and  $\beta$  are defined to be

$$\beta = \frac{\tau_{min}}{\tau_{min} - \mu_{min} r^{(1-a)}/\omega} \quad \text{and} \quad \alpha = \frac{\omega}{\beta \tau_{min}^{\beta-1} r^{(1-a)}}. \quad (\text{B.27})$$

These choices for  $\alpha$  and  $\beta$  ensure that the profile functions and their first derivatives are continuous. We use the following values for the parameters

$$\begin{aligned} \tau_{min} &= 2\mu_{min} r^{1-a}/\omega \\ \mu_{min} &= 0.3 \text{ GeV}. \end{aligned} \quad (\text{B.28})$$

We define our scale variations via

$$\begin{aligned} \epsilon_{S/J} = 1/2 &\rightarrow +50\% \text{ variation,} \\ \epsilon_{S/J} = -1/2 &\rightarrow -50\% \text{ variation,} \\ \epsilon_{S/J} = 0 &\rightarrow \text{Canonical scale,} \end{aligned}$$

and take the final scale variation bands as the envelope of the set of bands from the individual variations.

## B.2 MATCHING COEFFICIENTS AND CONSISTENCY CHECKS

### B.2.1 Evaluation of matching coefficients

In pure dimensional regularization all diagrams contributing to the FFs vanish, and the only diagrams that contribute to the angularity FJF for quarks are Figs. 3a) and 3b) of Ref. [54]. For Fig. 3a) we get

$$\frac{C_F \alpha_s}{2\pi} \frac{(4\pi\mu^2)^\epsilon (1-\epsilon)}{\Gamma[1-\epsilon]} \frac{1-z}{1-a/2} \omega^{2a\epsilon/(2-a)} (1-z)^{-2(1-a)\epsilon/(2-a)} \times \left(1 + \frac{(1-z)^{1-a}}{z^{1-a}}\right)^{2\epsilon/(2-a)} \frac{1}{s_a^{1+2\epsilon/(2-a)}}, \quad (\text{B.29})$$

and for Fig. 3b) we get

$$\frac{C_F \alpha_s}{2\pi} \frac{2z}{1-a/2} \frac{(4\pi\mu^2)^\epsilon}{\Gamma[1-\epsilon]} \omega^{2a\epsilon/(2-a)} \frac{1}{(1-z)^{1+2(1-a)\epsilon/(2-a)}} \times \left(1 + \frac{(1-z)^{1-a}}{z^{1-a}}\right)^{2\epsilon/(2-a)} \frac{1}{s_a^{1+2\epsilon/(2-a)}}, \quad (\text{B.30})$$

where  $s_a = \omega^2 \tau_a$ . The first expression is singular as  $\tau_a \rightarrow 0$  the second is singular as  $z \rightarrow 1$  and  $\tau_a \rightarrow 0$ , but the singularities are regulated by dimensional regularization. Employing the distributional identity

$$\frac{1}{(1-z)^{1+\epsilon}} = -\frac{1}{\epsilon} \delta(1-z) + \left(\frac{1}{1-z}\right)_+ - \epsilon \left(\frac{\ln(1-z)}{1-z}\right)_+ + \dots, \quad (\text{B.31})$$

and similarly for  $\tau_a$  we find for the divergent terms

$$\frac{C_F \alpha_s}{2\pi} \left( \delta(s_a) \delta(1-z) \left[ \frac{2-a}{1-a} \frac{1}{\epsilon^2} + \frac{2-a}{1-a} \frac{1}{\epsilon} \ln\left(\frac{\mu^2}{\omega^2}\right) + \frac{3}{2\epsilon} \right] - \frac{1}{1-a} \frac{2}{\epsilon} \delta(1-z) \frac{1}{\omega^2} \left[ \frac{1}{\tau_a} \right]_+ - \delta(s_a) \frac{1}{\epsilon} P_{qq}(z) \right), \quad (\text{B.32})$$

where  $P_{qq}$  is defined in Eq. (4.13). The first four terms in this expression are the expected UV poles for the angularity jet function (multiplied by  $\delta(1-z)$ ), see Eq. (3.37) of Ref. [146]. In order to simplify this expression we have redefined  $4\pi e^{-\gamma_E} \mu^2 \rightarrow \mu^2$ , i.e., we are working in the  $\overline{MS}$  scheme. The last term is the expected UV pole in the perturbative evaluation of the QCD fragmentation function. Since  $\mathcal{G}_i(\tau_a, z, \mu)$  is expected to evolve like the angularity



jet function, this is the correct structure of UV divergences implied by Eq. (4.2). The finite pieces are given by

$$\begin{aligned}
\frac{1}{\omega^2} \frac{\mathcal{J}_{qq}(\tau_a, z, \mu)}{2(2\pi)^3} &= \frac{C_F \alpha_s}{2\pi} \frac{1}{\omega^2} \left\{ \delta(\tau_a) \delta(1-z) \frac{2-a}{1-a} \left( -\frac{\pi^2}{12} + \frac{1}{2} \ln^2 \left( \frac{\mu^2}{\omega^2} \right) \right) \right. \\
&\quad + \delta(\tau_a) \left( 1-z - \left[ \ln \left( \frac{\mu^2}{\omega^2} \right) + \frac{1}{1-a/2} \ln \left( 1 + \frac{(1-z)^{1-a}}{z^{1-a}} \right) \right] \frac{1+z^2}{(1-z)_+} \right. \\
&\quad + \frac{1-a}{1-a/2} (1+z^2) \left( \frac{\ln(1-z)}{1-z} \right)_+ \Big) \\
&\quad + \left[ \frac{1}{\tau_a} \right]_+ \left( \frac{1}{1-a/2} \frac{1+z^2}{(1-z)_+} - \delta(1-z) \frac{2}{1-a} \ln \left( \frac{\mu^2}{\omega^2} \right) \right) \\
&\quad \left. + \frac{2\delta(1-z)}{(1-a)(1-a/2)} \left[ \frac{\ln \tau_a}{\tau_a} \right]_+ \right\}. \tag{B.33}
\end{aligned}$$

In the limit  $a \rightarrow 0$  this becomes

$$\begin{aligned}
\frac{1}{\omega^2} \frac{\mathcal{J}_{qq}(\tau_0, z, \mu)}{2(2\pi)^3} &= \frac{C_F \alpha_s}{2\pi} \left\{ \delta(s) \delta(1-z) \left( -\frac{\pi^2}{6} + \ln^2 \left( \frac{\mu^2}{\omega^2} \right) \right) \right. \\
&\quad + \delta(s) \left( 1-z - \ln \left( \frac{\mu^2}{\omega^2} \right) \frac{1+z^2}{(1-z)_+} + \ln z P_{qq}(z) + (1+z^2) \left( \frac{\ln(1-z)}{1-z} \right)_+ \right) \\
&\quad \left. + \frac{1}{\omega^2} \left[ \frac{1}{\tau_0} \right]_+ \left( \frac{1+z^2}{(1-z)_+} - 2\delta(1-z) \ln \left( \frac{\mu^2}{\omega^2} \right) \right) + 2\delta(1-z) \frac{1}{\omega^2} \left[ \frac{\ln \tau_0}{\tau_0} \right]_+ \right\}, \tag{B.34}
\end{aligned}$$

where we have used  $\delta(\tau_0)/\omega^2 = \delta(s)$ . Using the following distributional identities

$$\begin{aligned}
\frac{1}{\omega^2} \left[ \frac{1}{\tau_0} \right]_+ &= \frac{1}{\omega^2} \left[ \frac{\omega^2}{s} \right]_+ = \frac{1}{\mu^2} \left[ \frac{\mu^2}{s} \right]_+ + \ln \left( \frac{\mu^2}{\omega^2} \right) \delta(s), \\
\frac{1}{\omega^2} \left[ \frac{\ln \tau_0}{\tau_0} \right]_+ &= \frac{1}{\omega^2} \left[ \frac{\ln(s/\omega^2)}{s/\omega^2} \right]_+ = \frac{1}{\mu^2} \left[ \frac{\ln(s/\mu^2)}{s/\mu^2} \right]_+ + \frac{\ln(\mu^2/\omega^2)}{\mu^2} \left[ \frac{\mu^2}{s} \right]_+ + \frac{1}{2} \ln \left( \frac{\mu^2}{\omega^2} \right) \delta(s), \tag{B.35}
\end{aligned}$$

which are readily verified by integrating both sides over  $s$ , one finds that in the  $a \rightarrow 0$  limit the finite piece is given by

$$\begin{aligned}
\frac{\mathcal{J}_{qq}(s, z, \mu)}{2(2\pi)^3} &= \frac{C_F \alpha_s}{2\pi} \left\{ \delta(s) \left( 1-z + \ln z P_{qq}(z) + (1+z^2) \left( \frac{\ln(1-z)}{1-z} \right)_+ - \frac{\pi^2}{6} \delta(1-z) \right) \right. \\
&\quad \left. + \frac{1}{\mu^2} \left[ \frac{\mu^2}{s} \right]_+ \frac{1+z^2}{(1-z)_+} + 2\delta(1-z) \frac{1}{\mu^2} \left[ \frac{\ln(s/\mu^2)}{s/\mu^2} \right]_+ \right\}, \tag{B.36}
\end{aligned}$$

which agrees with the matching coefficient found in Eq. (2.32) of Ref. [54].

Next we calculate  $\mathcal{J}_{qg}(\tau_a, z, \mu)$ . Naively this is related to  $\mathcal{J}_{qg}(\tau_a, z, \mu)$  by the replacement  $z \rightarrow 1 - z$ . However, because in the convolution integral of Eq. (4.2) the argument of  $\mathcal{J}_{ij}(\tau_a, z/z', \mu)$  is never zero, there is no need to regulate poles of  $z$ . Therefore, a divergent factor of  $(1 - z)^{-1-\epsilon}$  in  $\mathcal{J}_{qg}(\tau_a, z, \mu)$  becomes in  $\mathcal{J}_{qg}(\tau_a, z, \mu)$

$$\frac{1}{z^{1+\epsilon}} = \frac{1}{z} - \epsilon \frac{\ln z}{z} + O(\epsilon^2). \quad (\text{B.37})$$

Thus,  $\mathcal{J}_{qg}(\tau_a, z, \mu)$  is obtained by making the substitution  $z \rightarrow 1 - z$  and then dropping all  $\delta(z)$  and plus prescriptions. This is true for the  $\mathcal{J}_{qg}(s, z, \mu)$  calculated in Ref. [54] and remains true for  $\mathcal{J}_{qg}(\tau_a, z, \mu)$ . We thus find for the divergent terms

$$\frac{1}{\omega^2} \frac{\mathcal{J}_{qg}^{div}(\tau_a, z, \mu)}{2(2\pi)^3} = -\frac{1}{\omega^2} \frac{C_F \alpha_s}{2\pi} \frac{1}{\epsilon} \delta(\tau_a) P_{gq}(z), \quad (\text{B.38})$$

where  $P_{gq}$  is given in Eq. (4.13). For the finite pieces we get

$$\begin{aligned} \frac{1}{\omega^2} \frac{\mathcal{J}_{qg}(\tau_a, z, \mu)}{2(2\pi)^3} = \frac{C_F \alpha_s}{2\pi} \frac{1}{\omega^2} \left\{ \delta(\tau_a) \left( z + \left[ \frac{1}{1-a/2} \ln \left( \frac{z^{1-a}(1-z)^{1-a}}{z^{1-a} + (1-z)^{1-a}} \right) \right. \right. \right. \\ \left. \left. \left. - \ln \left( \frac{\mu^2}{\omega^2} \right) \right] P_{gq}(z) \right) + \frac{1}{1-a/2} \left[ \frac{1}{\tau_a} \right]_+ P_{gq}(z) \right\}. \end{aligned} \quad (\text{B.39})$$

Again, these reproduce the matching coefficients of Ref. [54] in the  $a \rightarrow 0$  limit.

For the divergent contributions to  $\mathcal{J}_{gg}(\tau_a, z, \mu)$  we get (from the diagrams in Fig. 4 of Ref. [54])

$$\begin{aligned} \frac{1}{\omega^2} \frac{\mathcal{J}_{gg}^{div}(\tau_a, z, \mu)}{2(2\pi)^3} = \frac{C_A \alpha_s}{2\pi} \frac{1}{\omega^2} \left( \delta(\tau_a) \delta(1-z) \left[ \frac{2-a}{1-a} \frac{1}{\epsilon^2} + \frac{2-a}{1-a} \frac{1}{\epsilon} \ln \left( \frac{\mu^2}{\omega^2} \right) + \frac{\beta_0}{2C_A} \frac{1}{\epsilon} \right] \right. \\ \left. - \frac{1}{1-a} \frac{2}{\epsilon} \delta(1-z) \left[ \frac{1}{\tau_a} \right]_+ \right) - \frac{\alpha_s}{2\pi} \frac{1}{\omega^2} \delta(\tau_a) \frac{1}{\epsilon} \tilde{P}_{gg}(z), \end{aligned} \quad (\text{B.40})$$

where the  $\tilde{P}_{gg}(z)$  is the **full** QCD splitting function that includes the term proportional to  $\beta_0\delta(1-z)$ . For the finite parts of  $\mathcal{J}_{gg}(\tau_a, z, \mu)$  we find

$$\begin{aligned}
\frac{1}{\omega^2} \frac{\mathcal{J}_{gg}(\tau_a, z, \mu)}{2(2\pi)^3} = & \frac{C_A \alpha_s}{2\pi} \frac{1}{\omega^2} \left\{ \delta(\tau_a) \delta(1-z) \frac{2-a}{1-a} \left( -\frac{\pi^2}{12} + \frac{1}{2} \ln^2 \left( \frac{\mu^2}{\omega^2} \right) \right) \right. \\
& + \delta(\tau_a) \left( -P_{gg}(z) \left[ \ln \left( \frac{\mu^2}{\omega^2} \right) + \frac{1}{1-a/2} \ln \left( 1 + \frac{(1-z)^{1-a}}{z^{1-a}} \right) \right] \right. \\
& + \frac{1-a}{1-a/2} \frac{2(1-z+z^2)^2}{z} \left( \frac{\ln(1-z)}{1-z} \right)_+ \Bigg) \\
& + \left[ \frac{1}{\tau_a} \right]_+ \left( \frac{1}{1-a/2} P_{gg}(z) - \delta(1-z) \frac{2}{1-a} \ln \left( \frac{\mu^2}{\omega^2} \right) \right) \\
& \left. + \frac{2\delta(1-z)}{(1-a)(1-a/2)} \left[ \frac{\ln \tau_a}{\tau_a} \right]_+ \right\}, \tag{B.41}
\end{aligned}$$

where  $P_{gg}$  is given in Eq. (4.13). In the limit  $a \rightarrow 0$ , this expression reduces to  $\mathcal{J}_{gg}(s, z, \mu)/(16\pi^3)$  found in Eq. (2.33) of Ref. [54].

For the divergent contributions to  $\mathcal{J}_{gq}(\tau_a, z, \mu)$  we find

$$\frac{1}{\omega^2} \frac{\mathcal{J}_{gq}^{div}(\tau_a, z, \mu)}{2(2\pi)^3} = -\frac{1}{\omega^2} \frac{\alpha_s T_R}{2\pi} \frac{1}{\epsilon} \delta(\tau_a) P_{qg}(z). \tag{B.42}$$

For the finite parts we get

$$\begin{aligned}
\frac{1}{\omega^2} \frac{\mathcal{J}_{gq}(\tau_a, z, \mu)}{2(2\pi)^3} = & \frac{\alpha_s T_R}{2\pi} \frac{1}{\omega^2} \left\{ \frac{1}{1-a/2} \left[ \frac{1}{\tau_a} \right]_+ P_{qg}(z) + \delta(\tau_a) 2z(1-z) \right. \\
& \left. + \delta(\tau_a) P_{qg}(z) \left[ \frac{1}{1-a/2} \ln \left( \frac{z^{1-a}(1-z)^{1-a}}{z^{1-a} + (1-z)^{1-a}} \right) - \ln \left( \frac{\mu^2}{\omega^2} \right) \right] \right\}, \tag{B.43}
\end{aligned}$$

where  $P_{qg}$  is again given in Eq. (4.13). In the limit  $a \rightarrow 0$ , this expression reduces to  $\mathcal{J}_{gq}(s, z, \mu)/(16\pi^3)$  in Eq. (2.33) of Ref. [54].

### B.2.2 Sum Rules

The sum rules,

$$J_i(\tau_a) = \frac{1}{2(2\pi)^3} \sum_j \int_0^1 dz z \mathcal{J}_{ij}(\tau_a, z), \quad (\text{B.44})$$

can be checked for  $i = q$  by performing the integral

$$J_q(\tau_a) = \frac{1}{2(2\pi)^3} \sum_j \int_0^1 dz z \mathcal{J}_{qj}(\tau_a, z) \quad (\text{B.45})$$

$$= \frac{1}{2(2\pi)^3} \int_0^1 dz z (\mathcal{J}_{qq}(\tau_a, z) + \mathcal{J}_{qg}(\tau_a, z)) \quad (\text{B.46})$$

$$= \frac{1}{2(2\pi)^3} \int_0^1 dz z (\mathcal{J}_{qq}(\tau_a, z) + \mathcal{J}_{qq}(\tau_a, 1-z)) \quad (\text{B.47})$$

$$= \frac{1}{2(2\pi)^3} \int_0^1 dz \mathcal{J}_{qq}(\tau_a, z), \quad (\text{B.48})$$

where in the last line we changed variables to  $z \rightarrow 1-z$  in the 2nd term. Inserting the expression in Eq. (B.33) into this integral yields the  $J_q(\tau_a)$  found in Eq. (3.35) of Ref. [146].

In the case of the  $i = g$  we have

$$\begin{aligned} J_g(\tau_a) &= \frac{1}{2(2\pi)^3} \int_0^1 dz z (\mathcal{J}_{gg}(\tau_a, z) + \mathcal{J}_{gq}(\tau_a, z)) \\ &= \frac{1}{2(2\pi)^3} \int_0^1 dz \frac{\mathcal{J}_{gg}(\tau_a, z) + \mathcal{J}_{gq}(\tau_a, z)}{2}, \end{aligned} \quad (\text{B.49})$$

because both  $\mathcal{J}_{gg}(\tau_a, z)$  and  $\mathcal{J}_{gq}(\tau_a, z)$  are symmetric under  $z \rightarrow 1-z$ . The sum rule is easiest to verify by writing the  $d$ -dimensional expressions for  $\mathcal{J}_{gg}(\tau_a, z)$  and  $\mathcal{J}_{gq}(\tau_a, z)$  before expanding in  $\epsilon = (4-d)/2$ . We find

$$\begin{aligned} \frac{1}{\omega^2} \frac{\mathcal{J}_{gg}(\tau_a, z, \mu)}{2(2\pi)^3} &= \frac{1}{\omega^2} \left( \frac{4\pi\mu^2}{\omega^2} \right)^\epsilon \frac{C_A\alpha_s}{2\pi} \frac{1}{\Gamma[1-\epsilon]} \frac{1}{1-a/2} (z^{a-1} + (1-z)^{a-1})^{\frac{2\epsilon}{2-a}} \left( \frac{1}{\tau_a} \right)^{1+\frac{2\epsilon}{1-a}} \\ &\times \left( \frac{2z}{1-z} + \frac{2(1-z)}{z} + 2z(1-z) \right) \end{aligned} \quad (\text{B.50})$$

$$\begin{aligned} \frac{1}{\omega^2} \frac{\mathcal{J}_{gq}(\tau_a, z, \mu)}{2(2\pi)^3} &= \frac{1}{\omega^2} \left( \frac{4\pi\mu^2}{\omega^2} \right)^\epsilon \frac{T_R\alpha_s}{2\pi} \frac{1}{\Gamma[1-\epsilon]} \frac{1}{1-a/2} (z^{a-1} + (1-z)^{a-1})^{\frac{2\epsilon}{2-a}} \left( \frac{1}{\tau_a} \right)^{1+\frac{2\epsilon}{1-a}} \\ &\times \left( 1 - \frac{2}{1-\epsilon} z(1-z) \right). \end{aligned} \quad (\text{B.51})$$

Inserting these two expressions into Eq. (B.49) one obtains exactly the integral expression for the  $d$ -dimensional  $J_g(\tau_a)$  found in Eq. (4.22) of Ref. [46].

## APPENDIX C

### C.1 UNNORMALIZED AND NORMALIZED CROSS SECTIONS FOR BODWIN ET.AL

Fig. (36) shows the unnormalized (Eq. (4.28)) and normalized cross section (Eq. (4.33)) for Bodwin et al.'s LDME extractions [5]. The  $^3P_J^{[8]}$  channel contribution is negative, which is a feature of these LDMEs as it leads to a cancellation between the  $^3S_1^{[8]}$  and  $^3P_J^{[8]}$  channels, making the depolarizing  $^1S_0^{[8]}$  the dominant production channel of  $J/\psi$  for  $z > 0.5$ .

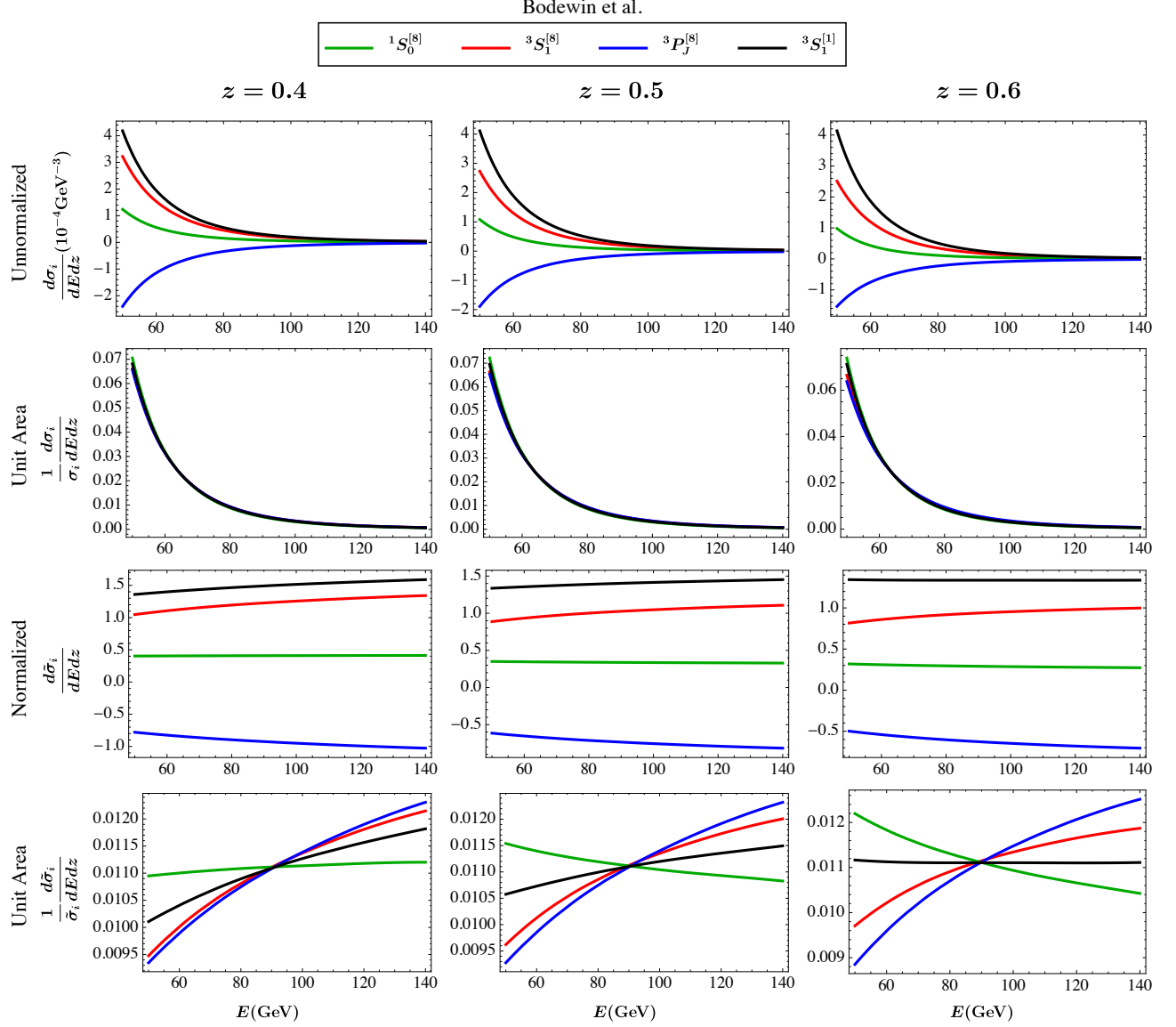


Figure 36: Unnormalized and normalized cross sections for Bodewin et al. extractions [5]. The conventions followed are same as in Fig. (30).

## C.2 UNNORMALIZED AND NORMALIZED CROSS SECTIONS FOR CHAO ET.AL

Fig. (37) shows the unnormalized (Eq. (4.28)) and normalized cross section (Eq. (4.33)) for Chao et al.'s LDME extractions [6]. Similar to Bodwin et al., these LDMEs result in a cancellation between the  $^3S_1^{[8]}$  and  $^3P_J^{[8]}$  channels.

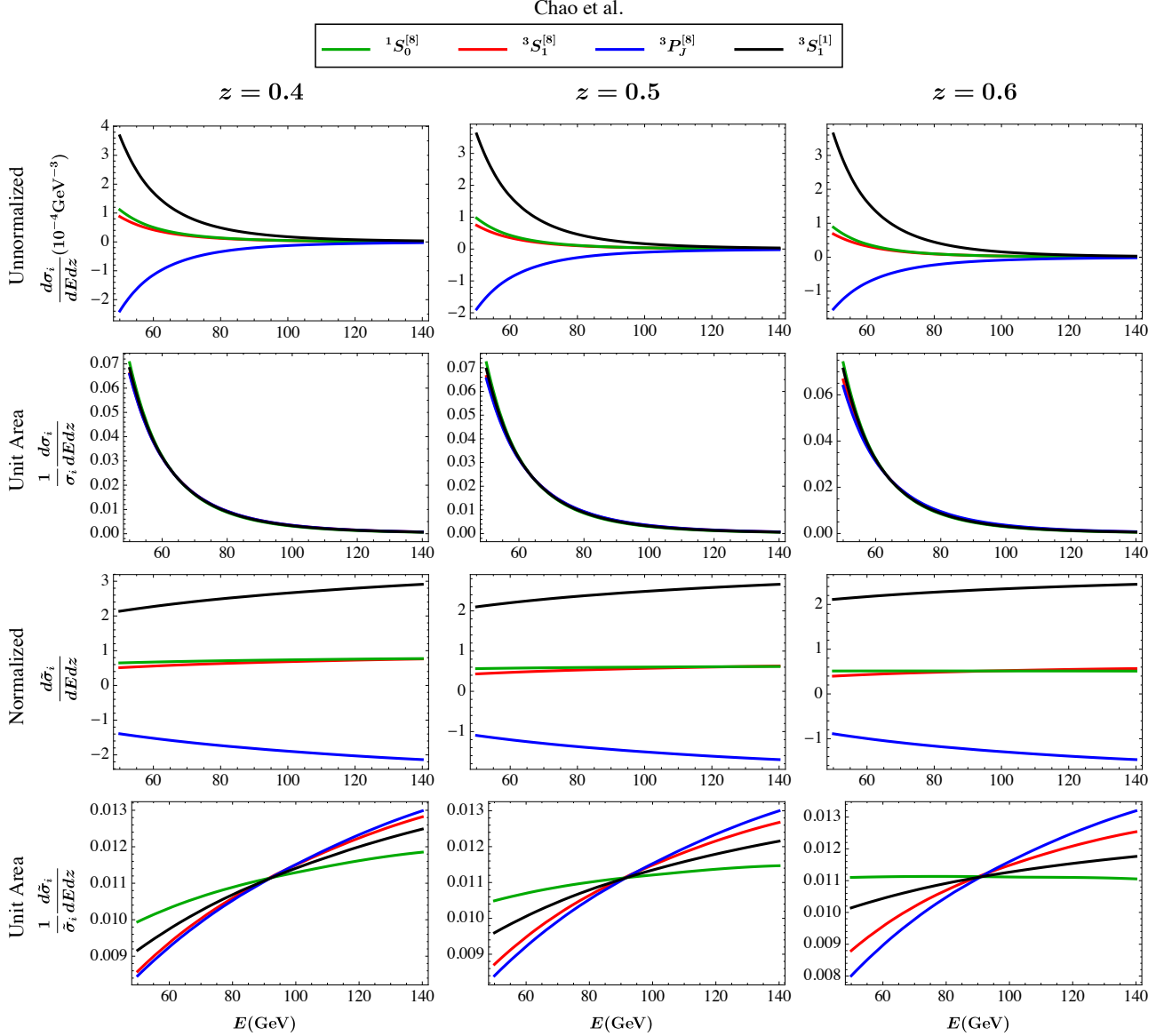


Figure 37: Unnormalized and normalized cross sections for Chao et al. extractions [6]. The conventions followed are same as in Fig. (30).

### C.3 NORMALIZATION USING ONLY COLOR OCTET CHANNELS

Fig. (38) shows the cross section for the different  $J/\psi$  production channels based on the LDMEs in Ref. [5] and Ref. [2, 3] with the contribution of  $^3S_1^{[1]}$  channel ignored in Eq. (4.33), i.e., setting  $\langle \mathcal{O}^{J/\psi}(^3S_1^{[1]}) \rangle$  to 0. Since  $^1S_0^{[8]}$  channel (green curves) has very different slopes for the two LDMEs, if the  $^1S_0^{[8]}$  channel dominates at high  $p_T$ , then one can distinguish between these two extractions. We don't include Chao et al.'s extractions [6] because it gives rise to a negative total cross section and so one can not ignore the color singlet contribution.

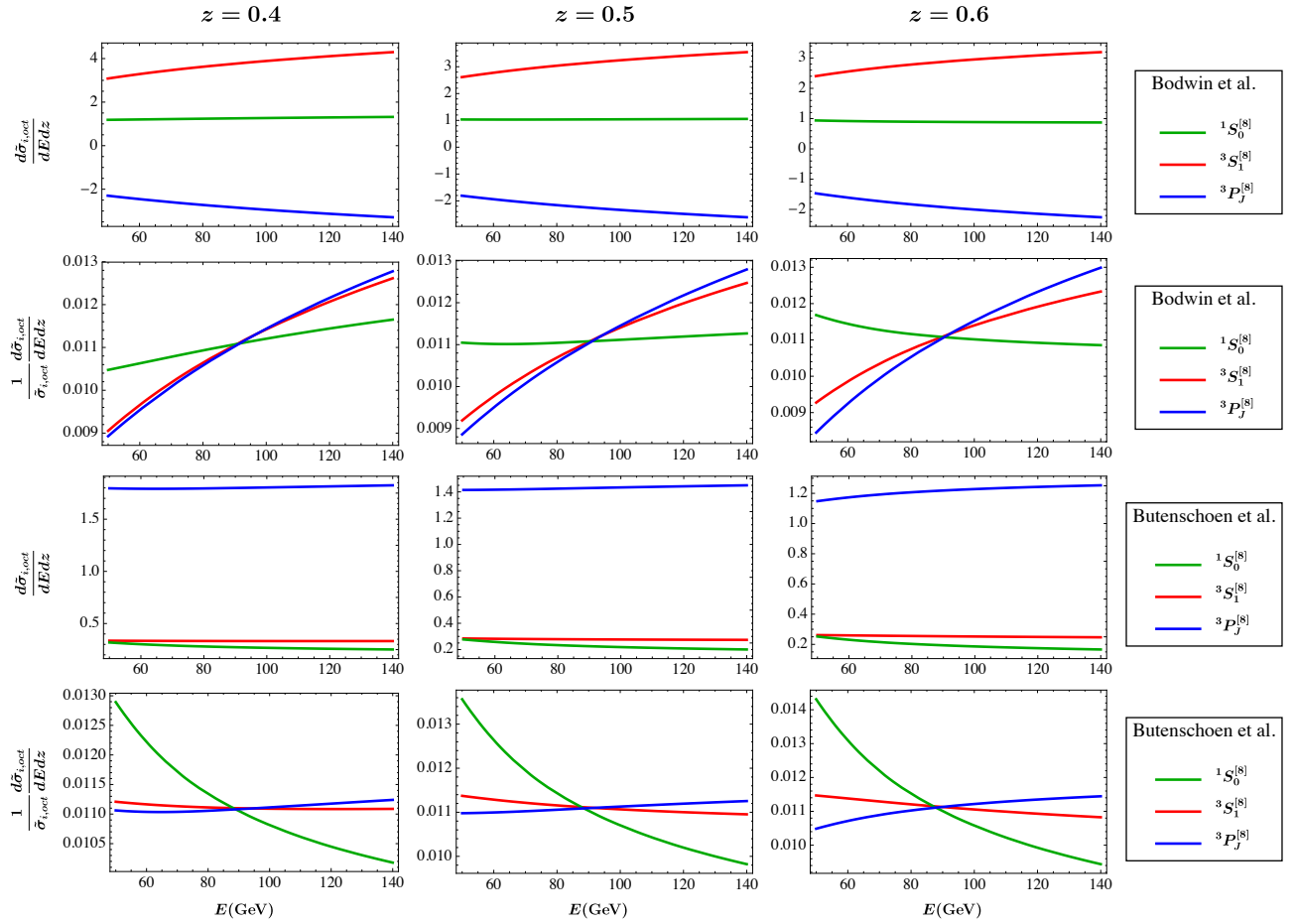
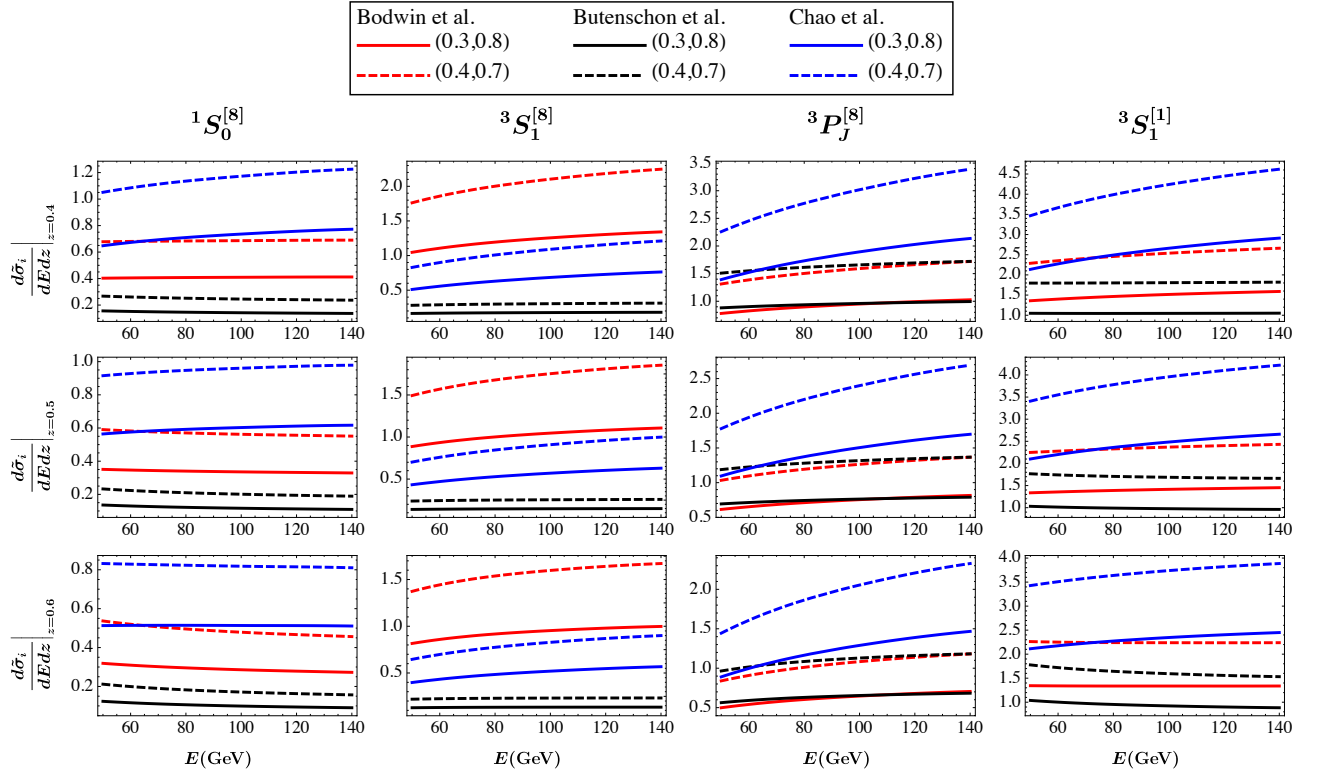


Figure 38: Cross section normalized by ignoring the  $^3S_1^{[1]}$  channel contribution in Eq. 4.33. The second and fourth row are obtained by normalizing the curves in the first and third row to unit area respectively.



## C.4 INSENSITIVITY TO $Z_{MIN}$ AND $Z_{MAX}$

Comparison of the normalized cross sections (Eq. (4.33)) for different values of  $z_{min}$  and  $z_{max}$  is shown. This confirms that the discussion in Subsection 4.3.3.1 is not sensitive to  $(z_{min}, z_{max})$  since the shapes of different LDMEs do not change. For validity of the factorization formula Eq. (4.28), we don't pick  $z_{min}$  too close to 0 and  $z_{max}$  too close to 1.



3

Figure 39: Solid curves correspond to  $(z_{min}, z_{max}) = (0.3, 0.8)$  and the dashed curves  $(z_{min}, z_{max}) = (0.4, 0.7)$ . Due to the change in normalization, all the curves shift upwards without changing their qualitative shapes.

## C.5 LOWER $z$ PLOTS

Fig. (40) shows the  $J/\psi$  production cross section (Eq. (4.35)) at lower  $z$  values for all the three LDME extractions [5, 2, 3, 6] used in this section.

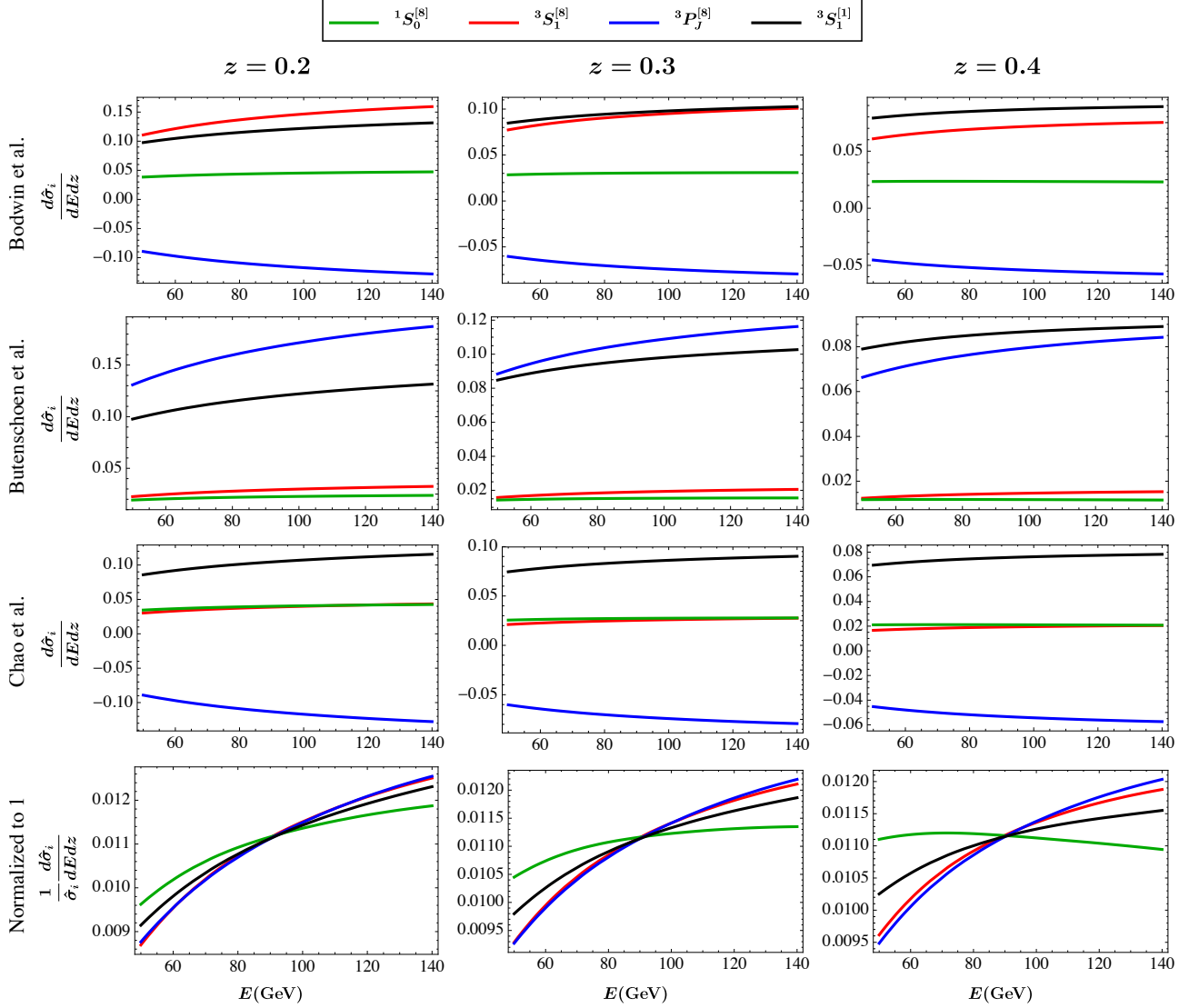


Figure 40: Lower  $z$  plots for the cross section (Eq. (4.35)). The conventions followed are same as those in Fig. (33).

## BIBLIOGRAPHY

- [1] E. Braaten, *Introduction to the NRQCD factorization approach to heavy quarkonium*, in *3rd International Workshop on Particle Physics Phenomenology Taipei, Taiwan, November 14-17, 1996*, 1996. [[hep-ph/9702225](#)].
- [2] M. Butenschoen and B. A. Kniehl, *World data of  $J/\psi$  production consolidate NRQCD at NLO*, *Phys. Rev.* **D84** (2011) 051501, [[1105.0820](#)].
- [3] M. Butenschoen and B. A. Kniehl, *Next-to-leading-order tests of NRQCD factorization with  $J/\psi$  yield and polarization*, *Mod. Phys. Lett.* **A28** (2013) 1350027, [[1212.2037](#)].
- [4] PARTICLE DATA GROUP collaboration, C. Patrignani et al., *Review of Particle Physics*, *Chin. Phys.* **C40** (2016) 100001.
- [5] G. T. Bodwin, H. S. Chung, U.-R. Kim and J. Lee, *Fragmentation contributions to  $J/\psi$  production at the Tevatron and the LHC*, *Phys. Rev. Lett.* **113** (2014) 022001, [[1403.3612](#)].
- [6] K.-T. Chao, Y.-Q. Ma, H.-S. Shao, K. Wang and Y.-J. Zhang,  *$J/\psi$  Polarization at Hadron Colliders in Nonrelativistic QCD*, *Phys. Rev. Lett.* **108** (2012) 242004, [[1201.2675](#)].
- [7] E. Rutherford, *The scattering of alpha and beta particles by matter and the structure of the atom*, *Phil. Mag. Ser.6* **21** (1911) 669–688.
- [8] J. Chadwick, *The Existence of a Neutron*, *Proc. Roy. Soc. Lond.* **A136** (1932) 692–708.
- [9] H. Yukawa, *On the Interaction of Elementary Particles I*, *Proc. Phys. Math. Soc. Jap.* **17** (1935) 48–57.
- [10] C.-N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, *Phys. Rev.* **96** (1954) 191–195.
- [11] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, *Phys. Rev. Lett.* **30** (1973) 1343–1346.
- [12] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, *Phys. Rev. Lett.* **30** (1973) 1346–1349.

- [13] G. F. Sterman and S. Weinberg, *Jets from Quantum Chromodynamics*, [\*Phys. Rev. Lett.\* \*\*39\*\* \(1977\) 1436](#).
- [14] G. P. Salam, *Towards Jetography*, [[0906.1833](#)].
- [15] N. Brambilla et al. [\*Eur. Phys. J.\* \*\*C71\*\* \(2011\) 1534](#), [[1010.5827](#)].
- [16] C. W. Bauer, S. Fleming and M. E. Luke, *Summing Sudakov logarithms in  $B \rightarrow X(s\gamma)$  in effective field theory*, [\*Phys. Rev.\* \*\*D63\*\* \(2000\) 014006](#), [[hep-ph/0005275](#)].
- [17] C. W. Bauer, S. Fleming, D. Pirjol and I. W. Stewart, *An Effective field theory for collinear and soft gluons: Heavy to light decays*, [\*Phys. Rev.\* \*\*D63\*\* \(2001\) 114020](#), [[hep-ph/0011336](#)].
- [18] C. W. Bauer and I. W. Stewart, *Invariant operators in collinear effective theory*, [\*Phys. Lett.\* \*\*B516\*\* \(2001\) 134–142](#), [[hep-ph/0107001](#)].
- [19] C. W. Bauer, D. Pirjol and I. W. Stewart, *Soft collinear factorization in effective field theory*, [\*Phys. Rev.\* \*\*D65\*\* \(2002\) 054022](#), [[hep-ph/0109045](#)].
- [20] C. W. Bauer, S. Fleming, D. Pirjol, I. Z. Rothstein and I. W. Stewart, *Hard scattering factorization from effective field theory*, [\*Phys. Rev.\* \*\*D66\*\* \(2002\) 014017](#), [[hep-ph/0202088](#)].
- [21] C. W. Bauer, D. Pirjol and I. W. Stewart, *Power counting in the Soft-Collinear Effective Theory*, [\*Phys. Rev.\* \*\*D66\*\* \(2002\) 054005](#), [[hep-ph/0205289](#)].
- [22] V. Chung, *Infrared Divergence in Quantum Electrodynamics*, [\*Phys. Rev.\* \*\*140\*\* \(1965\) B1110–B1122](#).
- [23] P. P. Kulish and L. D. Faddeev, *Asymptotic conditions and infrared divergences in quantum electrodynamics*, [\*Theor. Math. Phys.\* \*\*4\*\* \(1970\) 745](#).
- [24] R. Horan, M. Lavelle and D. McMullan, *Asymptotic dynamics in quantum field theory: When does the coupling switch off?*, 2000. [[hep-th/0002206](#)].
- [25] T. Becher, A. Broggio and A. Ferroglia, *Introduction to Soft-Collinear Effective Theory*, [\*Lect. Notes Phys.\* \*\*896\*\* \(2015\) pp.1–206](#), [[1410.1892](#)].
- [26] W. E. Caswell and G. P. Lepage [\*Phys. Lett.\* \*\*B167\*\* \(1986\) 437](#).
- [27] G. T. Bodwin, E. Braaten and G. P. Lepage, *Rigorous QCD analysis of inclusive annihilation and production of heavy quarkonium*, [\*Phys. Rev.\* \*\*D51\*\* \(1995\) 1125–1171](#), [[hep-ph/9407339](#)].
- [28] N. Brambilla, A. Pineda, J. Soto and A. Vairo, *Effective field theories for heavy quarkonium*, [\*Rev. Mod. Phys.\* \*\*77\*\* \(2005\) 1423](#), [[hep-ph/0410047](#)].

- [29] E. Braaten and Y.-Q. Chen, *Dimensional regularization in quarkonium calculations*, *Phys. Rev.* **D55** (1997) 2693–2707, [[hep-ph/9610401](#)].
- [30] Y.-Q. Ma, J.-W. Qiu and H. Zhang, *Heavy quarkonium fragmentation functions from a heavy quark pair. I. S wave*, *Phys. Rev.* **D89** (2014) 094029, [[1311.7078](#)].
- [31] L. Dai, C. Kim and A. K. Leibovich, *Fragmentation of a Jet with Small Radius*, *Phys. Rev.* **D94** (2016) 114023, [[1606.07411](#)].
- [32] L. Dai, C. Kim and A. K. Leibovich, *Fragmentation to a jet in the large  $z$  limit*, *Phys. Rev.* **D95** (2017) 074003, [[1701.05660](#)].
- [33] M. Dasgupta, F. Dreyer, G. P. Salam and G. Soyez, *Small-radius jets to all orders in QCD*, *JHEP* **04** (2015) 039, [[1411.5182](#)].
- [34] M. Dasgupta, F. A. Dreyer, G. P. Salam and G. Soyez, *Inclusive jet spectrum for small-radius jets*, *JHEP* **06** (2016) 057, [[1602.01110](#)].
- [35] B. Jager, M. Stratmann and W. Vogelsang, *Single inclusive jet production in polarized  $pp$  collisions at  $O(\alpha_s^3)$* , *Phys. Rev.* **D70** (2004) 034010, [[hep-ph/0404057](#)].
- [36] A. Mukherjee and W. Vogelsang, *Jet production in (un)polarized  $pp$  collisions: dependence on jet algorithm*, *Phys. Rev.* **D86** (2012) 094009, [[1209.1785](#)].
- [37] T. Becher, M. Neubert, L. Rothen and D. Y. Shao, *Effective Field Theory for Jet Processes*, *Phys. Rev. Lett.* **116** (2016) 192001, [[1508.06645](#)].
- [38] Y.-T. Chien, A. Hornig and C. Lee, *Soft-collinear mode for jet cross sections in soft collinear effective theory*, *Phys. Rev.* **D93** (2016) 014033, [[1509.04287](#)].
- [39] T. Becher, M. Neubert, L. Rothen and D. Y. Shao, *Factorization and Resummation for Jet Processes*, *JHEP* **11** (2016) 019, [[1605.02737](#)].
- [40] D. W. Kolodrubetz, P. Pietrulewicz, I. W. Stewart, F. J. Tackmann and W. J. Waalewijn, *Factorization for Jet Radius Logarithms in Jet Mass Spectra at the LHC*, *JHEP* **12** (2016) 054, [[1605.08038](#)].
- [41] Z.-B. Kang, F. Ringer and I. Vitev, *The semi-inclusive jet function in SCET and small radius resummation for inclusive jet production*, *JHEP* **10** (2016) 125, [[1606.06732](#)].
- [42] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, *Longitudinally invariant  $K_t$  clustering algorithms for hadron hadron collisions*, *Nucl. Phys.* **B406** (1993) 187–224.
- [43] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev.* **D48** (1993) 3160–3166, [[hep-ph/9305266](#)].
- [44] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001, [[hep-ph/9707323](#)].

- [45] M. Cacciari, G. P. Salam and G. Soyez, *The Anti- $k(t)$  jet clustering algorithm*, *JHEP* **04** (2008) 063, [[0802.1189](#)].
- [46] S. D. Ellis, C. K. Vermilion, J. R. Walsh, A. Hornig and C. Lee, *Jet Shapes and Jet Algorithms in SCET*, *JHEP* **11** (2010) 101, [[1001.0014](#)].
- [47] W. M.-Y. Cheung, M. Luke and S. Zuberi, *Phase Space and Jet Definitions in SCET*, *Phys. Rev.* **D80** (2009) 114021, [[0910.2479](#)].
- [48] J. Chay, C. Kim and I. Kim, *Factorization of the dijet cross section in electron-positron annihilation with jet algorithms*, *Phys. Rev.* **D92** (2015) 034012, [[1505.00121](#)].
- [49] A. V. Manohar and I. W. Stewart, *The Zero-Bin and Mode Factorization in Quantum Field Theory*, *Phys. Rev.* **D76** (2007) 074002, [[hep-ph/0605001](#)].
- [50] C. W. Bauer, F. J. Tackmann, J. R. Walsh and S. Zuberi, *Factorization and Resummation for Dijet Invariant Mass Spectra*, *Phys. Rev.* **D85** (2012) 074006, [[1106.6047](#)].
- [51] L. F. Abbott, *The Background Field Method Beyond One Loop*, *Nucl. Phys.* **B185** (1981) 189–203.
- [52] T. Kaufmann, A. Mukherjee and W. Vogelsang, *Hadron Fragmentation Inside Jets in Hadronic Collisions*, *Phys. Rev.* **D92** (2015) 054015, [[1506.01415](#)].
- [53] M. Procura and I. W. Stewart, *Quark Fragmentation within an Identified Jet*, *Phys. Rev.* **D81** (2010) 074009, [[0911.4980](#)].
- [54] A. Jain, M. Procura and W. J. Waalewijn, *Parton Fragmentation within an Identified Jet at NNLL*, *JHEP* **05** (2011) 035, [[1101.4953](#)].
- [55] M. Procura and W. J. Waalewijn, *Fragmentation in Jets: Cone and Threshold Effects*, *Phys. Rev.* **D85** (2012) 114041, [[1111.6605](#)].
- [56] Y.-T. Chien, Z.-B. Kang, F. Ringer, I. Vitev and H. Xing, *Jet fragmentation functions in proton-proton collisions using soft-collinear effective theory*, *JHEP* **05** (2016) 125, [[1512.06851](#)].
- [57] A. Idilbi and C. Kim, *Factorization of Jet Mass Distribution in the small  $R$  limit*, [[1606.05429](#)].
- [58] S. Catani and M. Grazzini, *Collinear factorization and splitting functions for next-to-next-to-leading order QCD calculations*, *Phys. Lett.* **B446** (1999) 143–152, [[hep-ph/9810389](#)].
- [59] J. C. Collins and D. E. Soper, *Parton Distribution and Decay Functions*, *Nucl. Phys.* **B194** (1982) 445–492.

- [60] A. Metz and A. Vossen, *Parton Fragmentation Functions*, *Prog. Part. Nucl. Phys.* **91** (2016) 136–202, [[1607.02521](#)].
- [61] G. P. Salam and G. Soyez, *A Practical Seedless Infrared-Safe Cone jet algorithm*, *JHEP* **05** (2007) 086, [[0704.0292](#)].
- [62] M. Baumgart, A. K. Leibovich, T. Mehen and I. Z. Rothstein, *Probing Quarkonium Production Mechanisms with Jet Substructure*, *JHEP* **11** (2014) 003, [[1406.2295](#)].
- [63] Z.-B. Kang, F. Ringer and I. Vitev, *Jet substructure using semi-inclusive jet functions in SCET*, *JHEP* **11** (2016) 155, [[1606.07063](#)].
- [64] R. Bain, Y. Makris and T. Mehen, *Transverse Momentum Dependent Fragmenting Jet Functions with Applications to Quarkonium Production*, *JHEP* **11** (2016) 144, [[1610.06508](#)].
- [65] D. Neill, I. Scimemi and W. J. Waalewijn, *Jet axes and universal transverse-momentum-dependent fragmentation*, *JHEP* **04** (2017) 020, [[1612.04817](#)].
- [66] M. Ritzmann and W. J. Waalewijn, *Fragmentation in Jets at NNLO*, *Phys. Rev.* **D90** (2014) 054029, [[1407.3272](#)].
- [67] M. Dasgupta and G. P. Salam, *Resummation of nonglobal QCD observables*, *Phys. Lett.* **B512** (2001) 323–330, [[hep-ph/0104277](#)].
- [68] A. Banfi, G. Marchesini and G. Smye, *Away from jet energy flow*, *JHEP* **08** (2002) 006, [[hep-ph/0206076](#)].
- [69] C. W. Bauer, D. Pirjol and I. W. Stewart, *On Power suppressed operators and gauge invariance in SCET*, *Phys. Rev.* **D68** (2003) 034021, [[hep-ph/0303156](#)].
- [70] J. Chay, C. Kim, Y. G. Kim and J.-P. Lee, *Soft Wilson lines in soft-collinear effective theory*, *Phys. Rev.* **D71** (2005) 056001, [[hep-ph/0412110](#)].
- [71] G. P. Korchemsky and A. V. Radyushkin, *Renormalization of the Wilson Loops Beyond the Leading Order*, *Nucl. Phys.* **B283** (1987) 342–364.
- [72] I. A. Korchemskaya and G. P. Korchemsky, *On lightlike Wilson loops*, *Phys. Lett.* **B287** (1992) 169–175.
- [73] M. Neubert, *Advanced predictions for moments of the anti- $B \rightarrow X(s)$  gamma photon spectrum*, *Phys. Rev.* **D72** (2005) 074025, [[hep-ph/0506245](#)].
- [74] T. Becher and M. Neubert, *Threshold resummation in momentum space from effective field theory*, *Phys. Rev. Lett.* **97** (2006) 082001, [[hep-ph/0605050](#)].
- [75] S. Caron-Huot, *Resummation of non-global logarithms and the BFKL equation*, [[1501.03754](#)].



- [76] A. J. Larkoski, I. Moult and D. Neill, *Non-Global Logarithms, Factorization, and the Soft Substructure of Jets*, *JHEP* **09** (2015) 143, [[1501.04596](#)].
- [77] D. Neill, *The Edge of Jets and Subleading Non-Global Logs*, [[1508.07568](#)].
- [78] A. J. Larkoski, I. Moult and D. Neill, *The Analytic Structure of Non-Global Logarithms: Convergence of the Dressed Gluon Expansion*, *JHEP* **11** (2016) 089, [[1609.04011](#)].
- [79] T. Becher, B. D. Pecjak and D. Y. Shao, *Factorization for the light-jet mass and hemisphere soft function*, *JHEP* **12** (2016) 018, [[1610.01608](#)].
- [80] D. de Florian, P. Hinderer, A. Mukherjee, F. Ringer and W. Vogelsang, *Approximate next-to-next-to-leading order corrections to hadronic jet production*, *Phys. Rev. Lett.* **112** (2014) 082001, [[1310.7192](#)].
- [81] A. Banfi, M. Dasgupta, K. Khelifa-Kerfa and S. Marzani, *Non-global logarithms and jet algorithms in high- $p_T$  jet shapes*, *JHEP* **08** (2010) 064, [[1004.3483](#)].
- [82] M. Dasgupta, K. Khelifa-Kerfa, S. Marzani and M. Spannowsky, *On jet mass distributions in  $Z$ +jet and dijet processes at the LHC*, *JHEP* **10** (2012) 126, [[1207.1640](#)].
- [83] R. Abbate, M. Fickinger, A. H. Hoang, V. Mateu and I. W. Stewart, *Thrust at  $N^3LL$  with Power Corrections and a Precision Global Fit for  $\alpha_s(m_Z)$* , *Phys.Rev.* **D83** (2011) 074021, [[1006.3080](#)].
- [84] M. Neubert, *Factorization analysis for the fragmentation functions of hadrons containing a heavy quark*, [[0706.2136](#)].
- [85] M. Fickinger, S. Fleming, C. Kim and E. Mereghetti, *Effective field theory approach to heavy quark fragmentation*, *JHEP* **11** (2016) 095, [[1606.07737](#)].
- [86] S. Fleming and O. Z. Labun, *Rapidity Divergences and Deep Inelastic Scattering in the Endpoint Region*, *Phys. Rev.* **D91** (2015) 094011, [[1210.1508](#)].
- [87] R. Bain, L. Dai, A. Hornig, A. K. Leibovich, Y. Makris and T. Mehen, *Analytic and Monte Carlo Studies of Jets with Heavy Mesons and Quarkonia*, *JHEP* **06** (2016) 121, [[1603.06981](#)].
- [88] R. Bain, L. Dai, A. K. Leibovich, Y. Makris and T. Mehen, *NRQCD Confronts LHCb Data on Quarkonium Production within Jets*, *Phys. Rev. Lett.* **119** (2017) 032002, [[1702.05525](#)].
- [89] L. Dai and P. Shrivastava, *Quarkonium Polarization and the Long Distance Matrix Elements Hierarchies using Jet Substructure*, *Phys. Rev.* **D96** (2017) 036020, [[1707.08629](#)].
- [90] L. G. Almeida, S. D. Ellis, C. Lee, G. Sterman, I. Sung and J. R. Walsh, *Comparing and counting logs in direct and effective methods of QCD resummation*, *JHEP* **04** (2014) 174, [[1401.4460](#)].



- [91] X. Liu, *SCET approach to top quark decay*, *Phys.Lett.* **B699** (2011) 87–92, [[1011.3872](#)].
- [92] A. Jain, M. Procura and W. J. Waalewijn, *Fully-Unintegrated Parton Distribution and Fragmentation Functions at Perturbative  $k_T$* , *JHEP* **1204** (2012) 132, [[1110.0839](#)].
- [93] A. Jain, M. Procura, B. Shotwell and W. J. Waalewijn, *Fragmentation with a Cut on Thrust: Predictions for B-factories*, *Phys. Rev.* **D87** (2013) 074013, [[1207.4788](#)].
- [94] C. W. Bauer and E. Mereghetti, *Heavy Quark Fragmenting Jet Functions*, *JHEP* **04** (2014) 051, [[1312.5605](#)].
- [95] C. F. Berger, T. Kucs and G. Sterman, *Event shape / energy flow correlations*, *Phys. Rev.* **D68** (2003) 014012, [[hep-ph/0303051](#)].
- [96] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer et al., *The Automated Computation of Tree-Level and Next-to-Leading Order Differential Cross Sections, and Their Matching to Parton Shower Simulations*, *JHEP* **07** (2014) 079, [[1405.0301](#)].
- [97] T. Sjostrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026, [[hep-ph/0603175](#)].
- [98] *An Introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159–177, [[1410.3012](#)].
- [99] M. Bahr et al., *Herwig++ Physics and Manual*, *Eur. Phys. J.* **C58** (2008) 639–707, [[0803.0883](#)].
- [100] V. G. Kartvelishvili and A. K. Likhoded, *Heavy Quark Fragmentation Into Mesons and Baryons*, *Sov. J. Nucl. Phys.* **29** (1979) 390.
- [101] B. A. Kniehl, G. Kramer, I. Schienbein and H. Spiesberger, *Finite-mass effects on inclusive B meson hadroproduction*, *Phys. Rev.* **D77** (2008) 014011, [[0705.4392](#)].
- [102] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J.* **C72** (2012) 1896, [[1111.6097](#)].
- [103] Z. Ligeti, I. W. Stewart and F. J. Tackmann, *Treating the b quark distribution function with reliable uncertainties*, *Phys. Rev.* **D78** (2008) 114014, [[0807.1926](#)].
- [104] A. Hornig, Y. Makris and T. Mehen, *Jet Shapes in Dijet Events at the LHC in SCET*, [[1601.01319](#)].
- [105] E. Braaten, K.-m. Cheung and T. C. Yuan, *Z0 decay into charmonium via charm quark fragmentation*, *Phys. Rev.* **D48** (1993) 4230–4235, [[hep-ph/9302307](#)].
- [106] E. Braaten and T. C. Yuan, *Gluon fragmentation into heavy quarkonium*, *Phys. Rev. Lett.* **71** (1993) 1673–1676, [[hep-ph/9303205](#)].

- [107] E. Braaten and S. Fleming, *Color octet fragmentation and the psi-prime surplus at the Tevatron*, *Phys. Rev. Lett.* **74** (1995) 3327–3330, [[hep-ph/9411365](#)].
- [108] E. Braaten and Y.-Q. Chen, *Helicity decomposition for inclusive  $J/\psi$  production*, *Phys. Rev.* **D54** (1996) 3216–3227, [[hep-ph/9604237](#)].
- [109] A. Buckley, J. Butterworth, L. Lonnblad, D. Grellscheid, H. Hoeth, J. Monk et al., *Rivet user manual*, *Comput. Phys. Commun.* **184** (2013) 2803–2819, [[1003.0694](#)].
- [110] C. Lee and G. Sterman, *Momentum flow correlations from event shapes: Factorized soft gluons and Soft-Collinear Effective Theory*, *Phys. Rev.* **D75** (2007) 014022, [[hep-ph/0611061](#)].
- [111] LHCb collaboration, R. Aaij et al., *Study of  $J/\psi$  production in jets*, [[1701.05116](#)].
- [112] LHCb collaboration, R. Aaij et al., *Measurement of the  $\eta_c(1S)$  production cross-section in proton-proton collisions via the decay  $\eta_c(1S) \rightarrow p\bar{p}$* , *Eur. Phys. J.* **C75** (2015) 311, [[1409.3612](#)].
- [113] M. Butenschoen, Z.-G. He and B. A. Kniehl,  *$\eta_c$  production at the LHC challenges nonrelativistic-QCD factorization*, *Phys. Rev. Lett.* **114** (2015) 092004, [[1411.5287](#)].
- [114] H. Han, Y.-Q. Ma, C. Meng, H.-S. Shao and K.-T. Chao,  *$\eta_c$  production at LHC and indications on the understanding of  $J/\psi$  production*, *Phys. Rev. Lett.* **114** (2015) 092005, [[1411.7350](#)].
- [115] H.-F. Zhang, Z. Sun, W.-L. Sang and R. Li, *Impact of  $\eta_c$  hadroproduction data on charmonium production and polarization within NRQCD framework*, *Phys. Rev. Lett.* **114** (2015) 092006, [[1412.0508](#)].
- [116] Z. Sun and H.-F. Zhang, *Reconciling charmonium production and polarization data within the nonrelativistic QCD framework*, [[1505.02675](#)].
- [117] B. Gong, L.-P. Wan, J.-X. Wang and H.-F. Zhang, *Polarization for Prompt  $J/\psi$  and  $\psi(2s)$  Production at the Tevatron and LHC*, *Phys. Rev. Lett.* **110** (2013) 042002, [[1205.6682](#)].
- [118] G. T. Bodwin, K.-T. Chao, H. S. Chung, U.-R. Kim, J. Lee and Y.-Q. Ma, *Fragmentation contributions to hadroproduction of prompt  $J/\psi$ ,  $\chi_{cJ}$ , and  $\psi(2S)$  states*, *Phys. Rev.* **D93** (2016) 034041, [[1509.07904](#)].
- [119] Z.-B. Kang, F. Ringer and I. Vitev, *Semi-inclusive jet cross sections within SCET*, *PoS QCDEV2016* (2017) 022, [[1609.07112](#)].
- [120] W. J. Waalewijn, *Calculating the Charge of a Jet*, *Phys. Rev.* **D86** (2012) 094030, [[1209.3019](#)].
- [121] G. T. Bodwin, *private communication*.

- [122] Z.-B. Kang, J.-W. Qiu, F. Ringer, H. Xing and H. Zhang,  *$j/\psi$  production and polarization within a jet*, [[1702.03287](#)].
- [123] P. Ilten, N. L. Rodd, J. Thaler and M. Williams, *Disentangling Heavy Flavor at Colliders*, [[1702.02947](#)].
- [124] Z. B. Kang, J. W. Qiu, F. Ringer, H. Xing and H. Zhang, *Phys. Rev. Lett.* **119** (2017) 032001, [[1702.03287](#)].
- [125] I. Belyaev, A. V. Berezhnoy, A. K. Likhoded and A. V. Luchinsky, *Comments on 'Study of  $J/\psi$  production in jets'*, *Mod. Phys. Lett.* **A32** (2017) 1771002, [[1703.09081](#)].
- [126] G. T. Bodwin, E. Braaten and G. P. Lepage, *Phys. Rev.* **D51** (1995) 1125, [[hep-ph/9407339](#)].
- [127] N. Brambilla, A. Pineda, J. Soto and A. Vairo, *Nucl.Phys.* **B566** (2000) 275, [[hep-ph/9907240](#)].
- [128] M. E. Luke, A. V. Manohar and I. Z. Rothstein, *Phys.Rev.* **D61** (2000) 074025, [[hep-ph/9910209](#)].
- [129] P. L. Cho and A. K. Leibovich, *Phys. Rev.* **D53** (1996) 150.
- [130] S. Fleming, A. K. Leibovich and I. Z. Rothstein, *Power counting and effective field theory for charmonium*, *Phys.Rev.* **D64** (2001) 036002, [[hep-ph/0012062](#)].
- [131] P. Cho and M. B. Wise, *Phys.Lett.* **B346** (1995) 129–136, [[hep-ph/9411303](#)].
- [132] M. Beneke and I. Z. Rothstein, *Psi-prime polarization as a test of color octet quarkonium production*, *Phys. Lett.* **B372** (1996) 157–164, [[hep-ph/9509375](#)].
- [133] M. Beneke and I. Z. Rothstein, *Hadro-production of quarkonia in fixed target experiments*, *Phys.Rev.* **D54** (1996) 2005, [[hep-ph/9603400](#)].
- [134] CDF collaboration, A. Abulencia et al., *Polarization of  $J/\psi$  and  $\psi(2S)$  mesons produced in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$ -TeV*, *Phys. Rev. Lett.* **99** (2007) 132001, [[0704.0638](#)].
- [135] CMS collaboration, S. Chatrchyan et al., *Measurement of the prompt  $J/\psi$  and  $\psi(2S)$  polarizations in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*, *Phys. Lett.* **B727** (2013) 381–402, [[1307.6070](#)].
- [136] LHCb collaboration, R. Aaij et al., *Measurement of  $J/\psi$  polarization in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*, *Eur. Phys. J.* **C73** (2013) 2631, [[1307.6379](#)].
- [137] I. W. Stewart, F. J. Tackmann and W. J. Waalewijn,  *$N$ -Jettiness: An Inclusive Event Shape to Veto Jets*, *Phys. Rev. Lett.* **105** (2010) 092002, [[1004.2489](#)].
- [138] S. D. Ellis, A. Hornig, C. Lee, C. K. Vermilion and J. R. Walsh, *JHEP* **1011** (2010) 101, [[1001.0014](#)].

- [139] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and Collider Physics*. Cambridge University Press, December, 2003.
- [140] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Eur.Phys.J.* **C63** (2009) 189–285, [[0901.0002](#)].
- [141] J. Campbell, F. Maltoni and F. Tramontano, *Phys. Rev. Lett.* **98** (2007) 252002.
- [142] B. Gong and J.-X. Wang, *Phys. Rev. Lett.* **100** (2008) 232001, [[0802.3727](#)].
- [143] B. Gong, X. Q. Li and J.-X. Wang, *Phys. Lett.* **B673** (2009) 197.
- [144] J. Lansberg, *Phys. Lett.* **B695** (2011) 149, [[1003.4319](#)].
- [145] M. Butenschoen and B. A. Kniehl, *Reconciling  $J/\psi$  Production at HERA, RHIC, Tevatron, and LHC with NRQCD Factorization at Next-to-Leading Order*, *Phys. Rev. Lett.* **106** (2011) 022003, [[1009.5662](#)].
- [146] A. Hornig, C. Lee and G. Ovanessian, *Effective predictions of event shapes: factorized, resummed and gapped angularity distributions*, *JHEP* **05** (2009) 122, [[0901.3780](#)].